

VERGLEICH VON SPRECHSTIMMEN MITTELS
DER SONAGRAPHIE

MAGDALENA LUDEWIG

1971



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41553801_0036

AUS DER KLINIK FÜR HALS-NASEN-OHREN-HEILKUNDE
DER UNIVERSITÄT MÜNCHEN

Direktor: Prof. Dr. H.H. Naumann

VERGLEICH VON SPRECHSTIMMEN MITTELS DER SONOGRAPHIE

Inaugural-Dissertation
zur Erlangung der Doktorwürde in der Zahnheilkunde
verfaßt und einer
Hohen Medizinischen Fakultät
der Ludwig-Maximilian-Universität
zu München

vorgelegt von
M a g d a l e n a L u d e w i g
aus München

München, den 28.1.1971

Gedruckt mit Genehmigung der
Medizinischen Fakultät
der Universität München

Referent: Prof. Dr. E. Loebell

Korreferent: Prof. Dr. H.H. Naumann

Dekan: Prof. Dr. W. Spann

Tag der mündlichen Prüfung: 8.6.1971



INHALTSVERZEICHNIS

	Seite
A EINLEITUNG	
I. Angewandte Sonographie in der Kriminalistik	4
II. Testung der Sonographie auf Brauchbarkeit zur Beweisführung	4
B LITERATUR	
I. Methoden der Stimmanalyse	5
1) Tonfrequenzspektrometrie	5
2) Suchtonanalyse	5
a) Sonographie	6
a1) Streifensonogramme	6
a2) Reliefdiagramme	8
II. Anwendung der Stimmanalyse	10
1) Forschung	10
a) Phonetik und Sprache	10
b) Zoologie	13
c) Geologie	14
2) Medizin	14
a) Spracherziehung Taubstummer	14
b) Beseitigung von Sprach- und Phonationsfehlern	15
c) Lärm- und Geräuschuntersuchung	15
d) Herztonaufzeichnung	15
3) Kriminalistik	15
a) Identifizierungshilfe	15
4) Zukünftige Anwendungsmöglichkeiten	16
III. Anatomie und Physiologie der Stimmbildung	17
1) Anatomie des Kehlkopfes und der Resonanzräume	17
2) Theorie der Stimmbildung	18
a) Physiologie	19
b) Physikalisches	20

	Seite
IV. Möglichkeiten zur Verstellung der Sprechstimme	22
1) Physiologische Verstellmöglichkeiten	22
2) Künstliche Verstellmöglichkeiten	23
V. Verfahren zur Identifizierung von Sprechern aus der Literatur	23
 C EIGENE UNTERSUCHUNGEN	
 I. Material und Versuchsbeschreibung	25
1) Die sonagraphischen Aufnahmen.	25
a) Der Sonagraph	25
b) Die Aufnahmetechnik	26
2) Die Aufnahmen mit dem Kathodenstrahloszillographen	26
a) Der Kathodenstrahloszillograph	26
b) Die Aufnahmetechnik	27
 II. Angaben über die Versuchspersonen	27
1) Anzahl	27
2) Alter	27
3) Geschlecht	27
4) Sprachzugehörigkeit	27
5) Sprachgeschulte Testpersonen	28
6) Ungeschulte Sprecher	28
 III. Wortwahl und Programm	28
 IV. Methodik	29
1) Veränderung der Tonhöhe	29
2) Tarnung	29
3) Fremdkörperwirkung	29
4) Veränderung der Mundart	30
 V. Auswertung	30
1) Fragestellung	30
2) Auswertung der gewonnenen Analysen	30

a)	Subjektiver visueller Bildvergleich für Streifen- und Reliefdiagramme	30
a1)	Ergebnisse	31
a2)	Statistischer Schluß	31
b)	Vergleichende Vermessung der Resonanzstreifen mit statistischer Prüfung für	31
b1)	gleiche Sprecher mit verschiedenen Sprechproben	31
b2)	verschiedene Sprecher mit unverstellten Sprechproben	33
c)	Vergleichende Vermessung der durch Stift zwischen den Zähnen oder durch Mundart verstellten Sprechweisen von "Apotheke"	34
d)	Amplituden- und Lautdauergestaltung	35
e)	Vergleichende Vermessung der Reliefdiagramme	36
e1)	Frequenzmessung mit dem Schreibstift des Sonagraphen und statistischer Schluß	36
e2)	Planimetrische Messungen	37
f)	Vergleich der Oszillographenbilder	37
f1)	Messung der Amplitudenmaxima	37
D	DISKUSSION	
I.	Diskussion der Ergebnisse	38
II.	Ergebnis für die Fragestellung	41
E	ZUSAMMENFASSUNG	41
F	ANHANG	44
I.	Außerliteratistische Quellen	44
II.	Abbildungen	44
III.	Graphische Darstellungen	52
IV.	Danksagungen	69
V.	Lebenslauf	70
G	LITERATURVERZEICHNIS	71

A EINLEITUNG

I. Angewandte Sonographie in der Kriminalistik

Sonographie bedeutet Aufzeichnung von Schallereignissen. Im Laufe der letzten 30 Jahre ist sie bis auf ihren heutigen Stand entwickelt worden und bietet auch jetzt noch genug Möglichkeiten zu Verbesserungen. Gelegentlich wird sie heute in der Kriminalistik benutzt, um Personen auf Grund von Tonbandaufnahmen zu identifizieren. Dabei handelt es sich hauptsächlich um die immer größer werdende Zahl von anonymen Anrufern, deren Anliegen von scherzhaft alkoholisierten Störungen über anonyme Beleidigungen, erotische Belästigungen, Drohungen, schwere Erpressungen bis zum erpresserischen Kinderraub reichen (H. HABERSBRUNNER).

Die Notwendigkeit, jeden Anruf über irgendeinen Fernsprecher ohne Nebengeräusch und Verzerrungen deutlich mitzuschneiden, birgt schon das erste große Problem dieser Methode. Weiterhin müßte bei einem späteren Vergleich mit verdächtigen Personen der Probetext über das gleiche Fernsprechgerät gesprochen werden wie die Tatstimme aufgenommen wird. Diese Bedingungen sind in der Regel nicht erfüllbar.

II. Testung der Sonographie auf Brauchbarkeit zur Beweisführung

Abgesehen von den technischen Schwierigkeiten soll untersucht werden, ob die eigentliche Stimmanalyse mit dem Sonographen das geeignete Mittel ist, die charakteristischen Momente einer Stimme und Sprechweise so aufzuzeichnen, daß sie sicher durch die vorhandenen Gemeinsamkeiten identifiziert werden kann. Wenn die sonographische Analyse als Indiz für eine Verurteilung ausreichen soll, muß hinreichend geprüft sein, ob die Stimme eines Menschen tatsächlich unveränderliche Kennzeichen wie Fingerabdrücke aufweist, und ob diese Methode geeignet ist, solche Merkmale aufzuzeigen.

· LITERATUR

I. Methoden der Stimmanalyse

Die Bemühungen um eine visuelle Darstellung der Sprache sind schon ziemlich alt. 1867 versucht MELVILLE BELL mit Hilfe handgemalter Symbole eine "visible speech" zu errichten. Die Bemühungen des Sohnes ALEXANDER BELL um eine Sicht- und Erfassbarmachung der Stimme und Sprache führen zur Erfindung des Telephons. Die Entwicklung eines "sound-soelectrograph" erfolgt 1941 für militärische Zwecke und bleibt geheim. Nach Kriegsende wird die Apparatur verbessert und allmählich auf den heutigen Stand entwickelt.

Die Analyse der Stimme und Sprache kann auf verschiedene Weise erfolgen: durch Tonfrequenzspektrometrie oder Suchtonanalyse.

1) Tonfrequenzspektrometrie

Unter diesem Schlagwort versteht man eine Analyse mit Hilfe von "band-pass"-Filtern, die kombinierte Spulen-Kondensatoren-Systeme darstellen und jeweils auf einen bestimmten Frequenzbereich eingestellt sind. Diese Methode ist in Verbindung mit einem Kathodenstrahloszillographen bekannt. Dabei werden die Informationen von den Filterausgängen, also die Partialfrequenzen, getrennt auf dem Bildschirm aufgezeigt. Die bei der Sprechprobe über ein dynamisches Mikrophon erzeugten Spannungen werden verstärkt und über Trennstufen gleichzeitig auf acht Filter geleitet. Das entstandene Frequenzspektrum schlüsselt die Probe in acht Intervalle auf, deren Größe von der Durchlässigkeit der Filter abhängig ist. Diese Meßwerte werden direkt in den Oszillographen geleitet, dort getrennt verstärkt und jeweils einem Komparator zugeführt. Dieser führt einen Spannungsvergleich zwischen der Meßwertspannung und einer Sinusspannung von ca. 20 kHz durch. Wenn beide Spannungen gleiches Potential haben, entsteht am Ausgang des Komparators ein Nadelimpuls, der, verstärkt im Impulsverstärker, zur Helltastung der Bildröhre verwendet wird. Die Lichtpunkte verschmelzen auf dem Bildschirm zu einem fortlaufenden Kurvenzug, wobei die Zeitablenkung horizontal verläuft und vertikal die Amplitudenauslenkung, also die Intensität, der Partialfrequenzen. Die Frequenzmuster müssen synchron mit einer Kamera aufgenommen werden, so daß sie als Negative oder Positive auszuwerten sind.

2) Suchtonalyse

Ihr Prinzip beruht auf der Differenzierung von Tonsignalen mit Hilfe einer hochfrequenten, veränderlichen Oszillatorfrequenz. In Abhängigkeit von der Größe der gerade zu analysierenden Frequenz "sucht" der Oszillator durch Abgabe einer passenden Frequenz an das Tongemisch und Differenzbildung beider die zu schreibende Frequenz aus der Mischung heraus. So entsteht ein lückenloses, kontinuierliches Frequenzspektrum.

a) Sonographie

Im Sonographen wird dieses Prinzip angewandt. Zusammengesetzte Schwingungen werden als Funktion von Frequenz, Intensität und Zeit analysiert. Im Sonogramm ist die Frequenz in vertikaler, die Zeit in horizontaler Richtung aufgetragen; die Intensität stellt sich im Grad der Schwärzung dar. Es kann ein Frequenzbereich von 85 - 8000 Hz kontinuierlich aufgezeichnet werden für einen Sprechvorgang von 2,4 sec. Dauer. Die zu analysierende Probe wird auf eine Magnettonscheibe aufgenommen, von dieser wieder abgenommen und in einem Wiedergabeverstärker verstärkt. Anschließend gelangt das Frequenzgemisch in einen Mischer, der von einem Zwischenfrequenzoszillator gespeist wird, welcher an die Schreibnadel des Sonographen gekoppelt ist. Diese windet sich an der routierenden Trommel spiralförmig von 85 bis zu 8000 Hz auf. Ihre jeweilige Stellung wird vom Oszillator registriert, und er gibt Frequenzen entsprechender Größe in den Mischer. Dort wird die Differenz der beiden Frequenzen gebildet und diese in den Intensitätsverstärker weitergeleitet. Dort befinden sich die Bandfilter, die den Frequenzbereich entweder eng (45 Hz) oder weit (300 Hz) an die Nadel weiterleiten. Je nach der Leistung der Teilfrequenzen entsteht an der Spitze des Glühdrahtes eine größere oder geringere Energie, die das hitzeempfindliche Registrierpapier dementsprechend unterschiedlich einbrennt. Das ergibt Dunkelheitsgrade von der hellgrauen Farbe des Papiers bis zu Tiefschwarz. Bei der Form der Frequenzdarstellung unterscheidet man hauptsächlich zwei Gruppen: die Darstellung der Resonanzen im Streifenmuster und die in Form von konturierten Feldern.

a1) Streifensonagramme

Diese können mit linearer oder logarithmischer Frequenzeinteilung her-

gestellt werden. Letztere hat den Vorteil, daß die Skala in den unteren, bedeutenderen Anteilen des Sonagramms eine feinere Messung zuläßt.

Die mehr oder weniger stark ausgeprägte horizontale Streifung entspricht bei allen Vokalen und vokalähnlichen Lauten den Resonanzen, die an bestimmten Frequenzen charakteristisch für jeden Vokal besonders verstärkt sind.

Die stimmlosen Konsonanten besitzen kein so regelmäßiges Klangspektrum, was auf ihre Bildungsweise als nicht-stimmhafte Laute zurückzuführen ist: Die Explosivlaute "p", "t" und "k" z.B. sind im Bild als Lücke dargestellt, bis am Ende des Lautes die Lippen plötzlich geöffnet werden. Dann entsteht durch die Reibungsmodulation eine schmale Zacke mit unregelmäßiger vertikaler Streifung. Sie wird "fill" genannt, da sie bestrebt ist, ein Segment des Musters auszufüllen. Bei den stimmlosen Verschlusslauten geht also die Verschlussmodulation der Reibungsmodulation voraus; der "fill" wird wegen seiner Schmalheit "spike fill" genannt.

Bei den stimmhaften Verschlusslauten kommt zu der im Prinzip gleichen Bildungsweise die Stimmbandschwingung dazu, die sich als einzelner Resonanzstreifen an der Grundlinie des Musters darstellt. Über dem Streifen liegt die Lücke, gefolgt vom "spike".

Die Zischlaute, bei denen nur Reibungsmodulation vorliegt, produzieren einen großen Frequenzbereich, der zufällig verteilt und unharmonisch zueinander ist.

Die verschiedenen Musterverschiebungen kommen durch die wechselseitigen Beeinflussungen der Buchstaben untereinander zustande. Bei der Lautfolge "a-p" z.B. liegt das Zentrum von "p" niedriger als das von "a", wodurch letzteres nach "p" heruntergebogen wird. Dabei gilt allgemein, daß die Bewegung der Artikulatoren vom Vokal zum Verschlusslaut oft schneller vor sich geht als umgekehrt. Dadurch wird der Einfluß des Verschlusslautes, der unmittelbar auf die Vokalstreifen folgt, nicht immer so offensichtlich, als wenn die Übertragungsbewegung dem Vokal vorangeht, oder wenn sie verlängert ist. Im ersten Fall erscheint im Muster gewöhnlich nur der "spike", im zweiten das volle Konsonantenbild. (POTTER, KOPP, GREEN)

a2) Reliefdiagramme

Eine andere Art von sonographischer Darstellung zeigt das Reliefdiagramm oder auch "contour display" genannt. Bei diesen Bildern ist die Lautprobe ebenfalls dreidimensional nach Frequenz, Zeit und Intensität aufgeschlüsselt. Die Resonanzen sind aber nicht in Streifen dargestellt, sondern in Flächenformen, die nach dem Prinzip topographischer Landkarten je nach Intensität in abgestuften Grautönen nach der Peripherie hin heller werden. Jedes Intensitätsfeld ist durch eine dunklere Hülllinie eingegrenzt.

Diese Methode wurde zur besseren Erfassung der Amplituden, d.h. des Energieinhaltes der aufgenommenen Äußerung, entwickelt. W. KOENIG und E. RUPPEL beschreiben verschiedene Verfahren dazu.

Die "dot-spacing-technique" bedient sich der Setzung von Punkten je nach Energie auf das Papier. Bei schwacher Intensität liegen sie weit auseinander und sind klein, bei großer rücken sie eng zusammen und sind groß dargestellt. Dieser Verfahrensweise ermangelt es aber an Detailgenauigkeit. Besser funktioniert die "contour-display-technique", wobei jede Linie des topographischen Bildes eine Teiltonintensität des Signals darstellt und die aufeinanderfolgenden Linien Steigerung der Intensität bedeuten. Der technische Vorgang stellt sich in Kürze so dar: Die Intensitätsschwankungen des Signals setzen einen Spiegel in entsprechende Vibrationen. Gebündeltes Licht, das den Spiegel trifft, wirft dieser durch eine Scheibe, die mit Schlitzen versehen ist, auf eine Photozelle. Die bei Lichtdurchtritt erzeugten Stromstöße bilden Punkte auf dem Papier, die genügend eng zusammenrücken, so daß eine Konturlinie entsteht. Jeder Schlitz formt eine eigene Kontur. Noch eine andere Art stellt die "light-chopper-method" dar, bei der das Licht vom Spiegel auf eine routierende Scheibe trifft, die transparente und opake Anteile in unterschiedlicher Anordnung enthält. Je nach Einteilung der Platte entstehen verschiedene Bilder mit Hell-Dunkel Effekt. Eine Kombination von Schlitzen und transparent-opaken Abschnitten ergibt ein gutes Diagramm, bei dem die verschiedenen Schattierungen nochmals von Konturlinien eingehüllt sind. Die heute gebräuchlichen Konturdarstellungen bedienen sich meist dieser Methode.

Die Aufzeichnung erfolgt im Sonographen genauso wie beim Streifendiagramm, nur wird hier die Energie ständig stufenweise bei jeder Umdrehung geändert.

Die Flächenformen jeder Intensität variieren stark von Vokal zu Vokal. Die Formanten sind als dunkelste Flächen über große Frequenzbereiche dargestellt und in ihrer Lage verhältnismäßig ausgedehnt. Ihre Position und Anzahl entspricht sich für gleiche Vokale bei verschiedenen Sprechern im Großen und Ganzen. Für jeden Vokal liegen deutlich zwei bis drei Intensitätszentren vor, wobei das frequenzniedrigste bei allen Vokalen am ausgeprägtesten dargestellt ist.

Es sind noch weitere Möglichkeiten der Sonagrammgestaltung versucht worden. POTTER gibt einen Typ an, wobei die Intensitätsschwankungen in verschiedenen Farben vorliegen. Es können Sonagramme durch eine größere Trommel auch zeitlich gedehnt werden zur besseren Analyse. Sie beinhalten aber die gleiche Information.

Eine große Hilfe für die Auswertung bietet die automatische Frequenzregistrierung während des Ablaufs des Sonagramms. Die Partialfrequenzwerte können entweder an einem Spezialeinschub auf drei Stellen genau abgelesen werden, oder es werden auf dem Papier direkt mit dem Ablaufvorgang Frequenzmarkierungen eingezeichnet, die, als Skala angeordnet, die Messungen sehr präzisieren.

Weitere Variationen der sonographischen Aufzeichnungen sind die Darstellung der Amplitudenkurve und das Sectionverfahren. Bei ersterem werden alle Intensitäten jedes Lautes summiert und mit einer logarithmischen Amplitudenskala über einen Bereich von 24 dB auf den oberen 38 mm des Registrierpapiers dargestellt. Dabei wird jedes Signal um 10 dB gedämpft, was die Reduzierung der Amplitudenfunktionshöhe um 10 mm zur Folge hat.

Beim Section-Programm kann an jedem beliebigen Zeitpunkt von 0,4 - 2,4 sec. die Intensität-Frequenzabhängigkeit dargestellt werden, wobei die Intensität horizontal verläuft und die Frequenz vertikal. Die Frequenzskala ist hierbei umgekehrt angebracht. Sie reicht von 4000 Hz in der Papiermitte bis zu 0 Hz am oberen Rand des Registrierpapiers.

Es ist auch möglich, die Frequenzskala zu dehnen, so daß auf dem gleichen Raum statt aller 8000 Hz nur die unteren 4000 Hz erscheinen.

II. Anwendung der Stimmanalyse

1) Forschung

Die Methode der Sonographie bewährt sich für Forschungszwecke verschiedener Art.

a) Phonetik und Sprache

Die gesamte Phonetikforschung kann durch die Analyse von Stimmen, Worten und vor allem einzelner Laute in ihrer Kenntnis der Zusammenhänge bestätigt und erweitert werden.

Besondere Beachtung wird dabei der Entstehungsweise und dem Aufbau der Vokale geschenkt.

Vokale sind Öffnungslaute mit dem wesentlichen Merkmal der resonatorisch erzeugten Klanggestaltung. (v. ESSEN)

Vokalklänge bestehen nach HERMANN aus periodischen Schwingungen. Bei der phonischen Konsonanten besteht diese Periodizität auch, sie hat jedoch Beimischungen von Schwingungen, die durch den Luftstrom an den verschiedenen Artikulationsstellen und Räumen entstehen. Die bei der Vokalklangbildung auftretenden, besonders verstärkten Partialtöne werden als Formanten bezeichnet. Diese "formen" den charakteristischen Vokalklang. Über das Wesen der Formanten sind die Autoren verschiedener Auffassung.

HELMHOLTZ vertritt die Ansicht der harmonischen, d.h. in ganzzahligem Verhältnis zum Grundton stehenden Teiltöne, die je nach Modus der Stimmbandschwingungen und der Form des Resonanzraumes verschiedenartig verstärkt werden.

Nach HERMANN werden die Teiltöne nicht im Kehlkopf, sondern durch den intermittierenden Luftstrom aus der Glottis erst im Ansatzraum hergestellt. Die sekundär entstandenen Töne sind meist nicht harmonisch zum Grundton, sondern stehen in einem irrationalen Verhältnis zu ihm.

C. STUMPF wiederum ist der Ansicht, daß Vokalklänge nur harmonische Teilfrequenzen enthalten. Dabei sind unabhängig von der Stimmfarbe der

Versuchspersonen ein oder zwei charakteristische Teiltongebiete, die sich über mehrere Teiltöne erstrecken, verstärkt.

Die verstärkten Obertöne sind die Stimmresonanzen, die im Sonogramm zur Darstellung kommen, wobei ihre Lage den Frequenzen entspricht, an denen die Stimmresonanzen vorkommen. (POTTER)

Nach v. ESSEN sind die Formanten in geringem Grad verschiebbar, d.h. sie rücken mit steigender oder fallender Grundtonfrequenz bei demselben Vokal etwas in die Höhe bzw. in die Tiefe. Ihre absolute Höhe ist trotzdem feststehend. So liegt z.B. der Formant von "a" immer bei e^2 - as^2 . Nur "u" hat einen wandernden Formanten, er ist mit dem Grundton identisch. V. ESSEN stellt weiter fest, daß ein Vokal, auf einen Grundton gesungen, der im Formantbereich oder höher liegt, unkenntlich wird. Grundsätzlich sind jedoch stimmhaft gesprochene Vokale von gesungenen hinsichtlich der Zusammensetzung der Teiltöne bei gleich deutlicher Aussprache nicht verschieden. Er stellt auch fest, daß jeder Formant ein Zentrum hat, d.h. einen Teilton, der als stärkster der Gruppe hervortritt.

POTTER, KOPP und GREEN können anhand sonographischer Untersuchungen diese Ansicht bestätigen. Sie sehen die Bedeutung des Zentrums, das sie besonders für den zweiten Streifen des Teiltonspektrums annehmen, darin, daß durch seine Lage die nachfolgenden oder vorhergehenden oder auch die übrigen Formanten des gleichen Lautes beeinflußt werden. Sie finden weiter, daß gleiche Laute von verschiedenen Sprechern meist ungefähr die gleiche Struktur haben. Das begründen sie damit, daß anatomisch ähnliche Verhältnisse des Stimmapparates und der Artikulatoren vorliegen, daß eine Einheitlichkeit der Sprache besteht, da sie während des Lernens vom Individuum imitiert wird, und daß zur allgemeinen Norm der Verständlichkeit Sprachgewohnheitsformen zur Ausschaltung extremer und behindernder Unterschiede übernommen werden. Auch zwischen männlichen und weiblichen Stimmustern finden sie keinen größeren Unterschied, da die Stimmqualität keinen großen Einfluß auf die relative Lage der Formanten zueinander und die Stellung des Zentrums hat. Variieren können lediglich Form und Intensität der Streifung.

KOCK, POTTER und PETERSON finden später, daß Kinderstimmen im Gesamtspektrum am höchsten liegen, Männer am tiefsten und Frauen dazwischen.

Zum Formantzentrum stellen POTTER, KOPP und GREEN noch fest, daß seine Po-

sition in den Vokalbildern stark wechselt, und zwar abfallend in der Reihenfolge i-e-a-o-u. Auch die vokalähnlichen Laute, zu denen sie j, w, l, m, n, ng, r zählen, haben unterschiedliche Zentren, ebenso Diphthonge. Auch in der Größe des Zwischenraumes von der Grundlinie bis zum ersten Resonanzstreifen äußert sich ein Unterschied zwischen Frontvokalen (i, e und offenes a) und Rückvokalen (o, u und geschlossenes a). SOKOLOWSKY findet mit Hilfe von Fourieranalysen, daß in der Dominanz gewisser Partialtöne zwischen offener und geschlossener Form des gleichen Vokals Unterschiede bestehen. Nur bei "a" ist kaum ein Unterschied erkennbar. Er kommt zu dem Schluß, daß offene Vokale ein deutliches Bestreben zeigen, sich der Indifferenzlage, also der Mittellage des Vokals "a" zu nähern. Bei "o" und "e" geschieht das direkt, bei "i" und "u" auf dem Umweg über "e" bzw. "o".

C.G.M. FANT erkennt anhand von Sonagrammen, daß das Intensitätsniveau von Formanten durch Artikulationswechsel nicht geändert werden kann. Er stellt aber einen Zusammenhang zwischen der Formantstellung in der Frequenzskala und dem Grad der Intensität fest. So kann ein Verstärken der Stimme das Formantniveau proportional zu seiner Frequenzstellung erhöhen und umgekehrt kann es ein Abschwächen erniedrigen.

Die häufigsten Fragen, die im Zusammenhang mit der sonographischen Phonetikforschung gestellt werden, sind die der Art und Weise der Beeinflussung von Konsonant- und Vokalmustern in der fließenden Rede, sowie die Frage nach den informationstragenden Bestandteilen der Sprache. Mit Hilfe von synthetischen Vokalen und Konsonanten, die mit den natürlichen fast identisch sind, werden von MENON, MILLER und FANT Versuche über die Wechselbeziehungen zwischen den Lauten der Sprache angestellt. Die synthetischen Laute haben für Forschungszwecke den großen Vorteil, daß Sprachparameter willkürlich eingesetzt werden können, und jede einzelne Wirkung gesondert registrierbar ist.

Recht häufig werden auch die Unterschiede bzw. Übereinstimmungen von Sprechmustern bei verschiedenen Sprechern untersucht. JOOS, POTTER, PETERSON und BARNAY kommen zu dem übereinstimmenden Ergebnis, daß die Merkmale eines Sprechers in gewissen Grenzen immer wiederkehren, die Unterschiede bei verschiedenen Sprechern aber statistisch signifikant sind. Dabei variieren hauptsächlich die Positionen der unteren zwei Formanten. JOOS zieht aus

seinen Untersuchungen die weitere Schlußfolgerung, daß die Beeinflussungen zwischen Konsonanten und Vokalen nicht mechanistisch zu erklären sind, sondern vom Zentralnervensystem gesteuert werden, das für bestimmte Konsonantkombinationen bestimmte Artikulationsweisen vorschreibt. Durch Überschneiden der Innervationswellen erklärt er die Tatsache des Ineinanderergreifens der Formantübergänge von Laut zu Laut.

b) Zoologie

Einen großen Anwendungsbereich für sonographische Registrierungen stellt auch das Gebiet der Zoologie dar. Vor allem die Ornithologie bedient sich gern des Sonagrammen, um die Gesänge, Triller und "Gespräche" von Vögeln aufzunehmen und zu analysieren.

Nach G. THIELCKE singen die meisten Vögel für das menschliche Ohr zu schnell, so daß Einzelheiten des Gesanges nicht erfaßt werden können. Die optische Auswertung ist dafür wesentlich geeigneter, wobei er besonders auf die gedehnte Zeitskala Wert legt, die die Bestimmung der Pausen vereinfacht. Die Sonographie ermöglicht auch die genaue Erforschung des Prinzips, nach dem die Klangspektren von verschiedenen Vögeln aufgebaut sind. Es wird untersucht, ob sich die Lautäußerungen aus Geräuschen oder reinen Tönen oder aus Kombinationen beider zusammensetzen, und es wird die Verteilung der Lautstärken festgestellt. Die Unterschiede zwischen den einzelnen Vogelarten können so gut dargestellt werden. Die Untersuchung von Unterschiedlichkeiten in den Spektrogrammen gleicher Vogelarten bei verschiedenen Lautäußerungen dient der Verhaltensforschung. THIELCKE stellt fest, daß sich akustisch ähnliche Gesänge von verschiedenen Vogelarten auch im Sonagramm als ähnlich aufgebaut erweisen. Er schließt daraus, daß bei großer Ähnlichkeit der Klangspektren von verschiedenen Arten eine gemeinsame oder voneinander abhängige Herkunft wahrscheinlich ist. Zu ähnlichen Forschungszwecken werden auch die Stimmen von Affen und anderen Tiergattungen analysiert.

O. TODT untersucht u.a. auch mit sonographischen Mitteln die Gesangsstrophen und Strophenelemente der Amsel. Er führt wörtlich aus: "Ein Vergleich zwischen der Lautäußerung der Amsel und der Sprache des Menschen scheint vorerst verfrüht. Wollte man ihn späterhin vornehmen, so könnte man auf Grund gewisser Ähnlichkeiten der zeitlichen Strukturierung formal

gegenüberstellen: Strophenelemente und Phoneme, Strophenelementgruppen und Worte, Strophen und ganze Sätze." Die Erforschung der Laut- und Sprachentwicklung würde somit einen nicht unwesentlichen Beitrag zur Evolutionstheorie liefern können.

c) Geologie

Sogar auf dem Gebiet der Geophysik findet der Sonograph Anwendung. Es werden Erdbeben und Explosionswellen aufgezeichnet und analysiert.

M. EWING u.a. untersuchen sonographisch die Wellen eines Erdbebens, das westlich der Hawaiiinseln stattfand und in 4050 km Entfernung, in Kalifornien aufgenommen wird. Sie erhalten klare und eindeutige Sonagramme, die eine genaue Analyse der körperlichen und oberflächlichen Wellen zulassen.

Vom gleichen Autor stammen Untersuchungen über die Ausbreitung von Explosionsschall im seichten Wasser. Dabei zeigt sich, daß die höheren Teiltöne mit steigender Entfernung schneller verschwinden. Solche Forschungen sind möglicherweise für militärische Zwecke von Nutzen.

2) Medizin

In dem im weitesten Sinne medizinischen Bereich gibt es ebenso einige Anwendungsmöglichkeiten für den Sonographen.

a) Spracherziehung Taubstummer

Wie von POTTER, KOPP und GREEN beschrieben, findet die Sonographie bedeutende Anwendung bei dem Versuch, Taubstumme und Gehörgeschädigte zu einer normalen Laut- und Sprachgestaltung zu führen. Dabei fällt dem Sonographen hauptsächlich eine Kontrollfunktion für Lernende und Lehrenden zu. Einzelne Lautmuster werden vom Normalhörenden vorgesprochen und aufgezeichnet. Das so entstandene Lautbild wird auf einem Bildschirm festgehalten. Die Schüler formen so lange den Laut nach, bis das Klangbild ihrer Aussprache mit dem der normalen weitgehend übereinstimmt. Das sonographische Bild bringt nicht mehr als das Ohr hören kann. So wird auf dem Umweg über das Visuelle in der Großhirnrinde der Taubstummen eine akustische Vorstellung des Lautes geschaffen, die es ihnen ermöglicht, in normaler Art und Weise die Stimme zu bilden und zu artikulieren.

b) Beseitigung von Sprach- und Phonationsfehlern

Für Sprachfehler und Gehörschäden gilt im Prinzip das Gleiche; der Sonograph übt auch hier die Kontrolle aus. Er stellt aber gleichzeitig ein psychologisches Hilfsmittel für den Lehrer dar, indem durch Verdeutlichung des Lernerfolges und der Behandlungsfortschritte dem Patienten ein Anreiz zur weiteren Mitarbeit gegeben wird. Zusätzlich lassen sich auf Grund der verschiedenen Darstellung im Sonogramm die einzelnen Stimm- und Sprachstörungen gegeneinander abgrenzen und besser diagnostizieren.

Bei Taubstummheit und Aphonie findet man eine generelle Verbreiterung der Resonanzbänder mit teilweiser Zerstörung der Formanten. Dysphonie kennzeichnet eine schwache Ausbildung der oberen Formanten, die Pause bei den Explosivlauten fällt durch Fortdauer der Resonanz weg, und die Struktur ist nicht identifizierbar. Bei Gaumenspalten fehlen die Obertöne, die Resonanzen überdauern, Explosiva fallen aus, und es wird keine Struktur erkennbar. In allen Fällen beeinflussen sich die Laute gegenseitig nicht. Ösophagussprache zeichnet sich, vor allem im "narrow"-Bild, durch besonders stark gebrochene und unregelmäßige Formanten aus. Eine Trennung der Tonelemente und Verlagerung der Formanten findet man bei Dyslalie, ein völlig verschwommenes, ineinandergeherdes Bild bei Dysphemie. Dysarthrie äußert sich durch Fehlen der oberen Resonanzen und einen starken Tremor, Aphasie stellt sich ähnlich dar, das Klangbild ist jedoch stärker verwischt. (R. POTTER).

G. BÖHME stellt fest, daß bei Hemiplegie, Diplegie, Tetraplegie, Atetose, Chorea, Choreathetose und Torsionsdystonie sonographisch eine Differenzierung in einen "Dämpfungstyp" und einen "Resonanztyp" vorgenommen werden kann. Ersterer zeigt eine Verstärkung der tiefen Frequenzen bei Fehlen oder angedeuteter Darstellung der oberen, letzterer bringt eine Verstärkung sämtlicher Frequenzen. Auch nach Kehlkopfteilresektionen führt er sonographische Kontrollen durch und stellt im Formantenaufbau eine gute Übereinstimmung mit der normalen Sprache fest. Im höheren Frequenzbereich ergeben sich aber zusätzliche Geräuschanteile.

c) Lärm- und Geräuschuntersuchung

Diese beziehen sich hauptsächlich auf die Schallintensität und werden bevorzugt mit dem Amplituden-Section-Verfahren durchgeführt. PETERSEN und

RAISBECK stellen auf diese Weise Messungen zur quantitativen Bestimmung von Lärm an und zeigen die Beziehung zwischen Lärm und reinen Tönen auf.

d) Herztoneaufzeichnung

Herzschallaufnahmen zu diagnostischen Zwecken werden seit einiger Zeit im Klinikum Rechts der Isar in München mit dem Sonagraphen versuchsweise gemacht. Die Auswertung liefert ein Computer unter besonderer Berücksichtigung der bevorzugten Frequenzen, des Abstandes und des Verhältnisses der Frequenzbänder zueinander, der harmonischen Schwingungen und der Messung von Länge, Breite und Energie.

3) Kriminalistik

a) Identifizierungshilfe

Sonographie als Hilfsmittel zur Identifizierung von Tätern über Stimm-analyse wird in Deutschland so gut wie nicht angewendet. In Amerika sind auf diesem Gebiet schon weitgehendere Versuche gemacht worden mit erscheinend recht gutem Erfolg. Der sonographischen Identifizierung von Sprechstimmen liegt die Annahme zu Grunde, daß gleiche Artikulationsweise und Stimmqualität gleiche optische Bilder ergeben, d.h. vergleichbare Äußerungen eines Sprechers anhand der Form und Verteilung von Resonanzen und Formanten ohne weiteres zu identifizieren sind. Selbst bei willkürlicher Veränderung verschiedener Parameter bleibt soviel unbewußte Identität vorhanden, daß sie im Bild sichtbar und erkennbar wird. Das gilt besonders für zusammenhängende Texte, wobei Sprachrhythmus, Phasenablauf, Diktion und Stimmhöhe sehr häufig in ihre normale Art zurückfallen, was auf dem Sonagramm gut festzuhalten ist. Besonders das Zurückgleiten auf den eigenen Indifferenzton ist bei Stimmverstellung sehr häufig zu beobachten (R. FÄHRMANN). Schwieriger wird die Untersuchung bei Einzelworten, aber das kommt in der Kriminalpraxis so gut wie nie vor.

4) Zukünftige Anwendungsmöglichkeiten

Eine Identifizierungshilfe im nicht kriminalistischen Sinn wird das Sonagraph-Computer-System in Zukunft einmal darstellen, das im Geldverkehr

die Unterschrift ersetzen soll. Die Stimmanalyse des gesprochenen Namens aller Kunden wird gespeichert, bei Bedarf neu analysiert und vom Computer verglichen. Diesem Verfahren wird eine größere Identitätssicherheit als dem Vergleichen der Unterschriften zugemessen.

Überhaupt bietet die Kombination des Analysators mit einem auswertenden Computer eine weit größere Sicherheit zum fehlerfreien Vergleich der Stimmdiagramme. Dieses Verfahren ist von M.R. WEISS bereits zur Bestimmung von Formantfrequenz, Formantamplitude, Spektrumintensität und Stimmanteil bei Sprechproben erfolgreich angewendet worden.

III. Anatomie und Physiologie der Stimmbildung

1) Anatomie der Kehlkopfes und der Resonanzräume

Der Kehlkopf stellt ein Knorpelgerüst dar, das durch Bänder und Gelenke zusammengehalten wird und den Stimmapparat enthält. Der eigentliche Stimmapparat besteht aus den Stimmlippen; sie begrenzen die beim Anblasen der Luft aus der Trachea in Form und Weite veränderliche Stimmritze. Die Umwandlung der Trachea als Anblaserohr zum Mundstück beginnt am oberen Rand des Ringknorpels. Dort löst sich ein elastischer Schlauch vom starren Knorpel und zieht, gegen das Lumen vorbauchend, als *conus elasticus* hinten an die *processus vocales* und die Unterseite des Stellknorpels; vorne an den Schildknorpel beidseitig direkt neben der Mittellinie. Mit dem verdickten oberen freien Rand bildet er als *ligamentum vocale* das Mundstück. Dieses ist durch die Verbindung mit den beweglichen Stellknorpeln und die veränderbare Muskulatur sehr plastisch. Die Art und Weise der Einstellung des Mundstückes ist entscheidend für die Tonbildung. Der Bewegungsapparat des *Cricothyroideus*verbandes sorgt für die äußere und innere Spannung der Stimmlippen. Dabei gehen Einstellung und Spannung Hand in Hand.

Der Mechanismus der Stimmlippenverlängerung funktioniert folgendermaßen: Der Schildknorpel ist durch die untere Zungenbeinmuskulatur fixiert. Durch Kontraktion des *musculus cricothyroideus* wird der Cricoidbogen aufwärts bewegt bis zur größten Annäherung an den Schildknorpel. Gleichzeitig wird die *lamina* mit den Stellknorpeln nach unten rückwärts abgekippt. So entsteht eine Vergrößerung der Distanz zu Thyreoid- und Aryknorpel im Kehlkopfinneren, was eine Verlängerung und Anspannung der Stimmlippen bedeutet.

Der Stimmritzenverschluß in Phonationsstellung bewirkt die gleichzeitige Kippung der Aryknorpel nach vorn, wobei sich die processus vocales und die medialen Vorderkanten der Aryknorpel berühren und auch den Verschluß der pars cartilaginea herbeiführen, der durch die Fältelung des lockeren Schleimhautüberzuges vollendet wird.

Das Führungsband für alle Gleit- und Scharnierbewegungen des Stellgelenks bildet das ligamentum cricoarytenoideum. Es führt als elastisches Band die Stellknorpel nach Aufhören der Muskelwirkung in die Ruhelage zurück.

Den eigentlichen Stimmapparat bildet der Spannungsapparat. Durch Kontraktion des äußeren Spannungsmuskels (m. cricothyreoideus) wird die Schwingungszahl der Stimmbänder mit verstärkter Spannung erhöht, und die Tonhöhe steigt. Allgemein bedeutet Verlängerung der Stimmbänder Tonerhöhung, Verkürzung Tonvertiefung. Der m. vocalis bestimmt durch isotonische Kontraktion die Stimmbandlänge mit, seine isometrische Kontraktion reguliert die Spannung.

Parallel zu den Stimmlippen liegen die Taschenfalten. Bei Inspiration ist ihr freier Rand gerade gestreckt, bei Phonation bildet er nach kaudal einen offenen Bogen und öffnet den Eingang in den ventriculus laryngis weit. Von dort wird bei Respiration ein schleimiges Sekret ausgepreßt, das die Stimmlippen immer feucht hält und die Voraussetzung für einen reinen Ton schafft. (BERENDES)

Der offene Ventrikel dient als Resonanzraum. Solche stellt auch das als "Ansatzrohr" bezeichnete System von Hohlräumen im Kopfbereich dar. Dazu gehören der Hypo- und Mesopharynx, der Epipharynx, die Nase mit ihren Nebenhöhlen, die Mundhöhle mit Zunge, Zähnen und mimischer Muskulatur, wobei besonders die Stellung der Lippen von Wichtigkeit ist. (LULLIES)

2) Theorie der Stimmbildung

Die Stimmbildung setzt sich aus einer physiologischen und einer physikalischen Komponente zusammen. Es gibt viele Theorien über die Stimmbildung, die im wesentlichen übereinstimmen. Ich beschränke mich hier auf die Darstellung der Tonerzeugung nach LULLIES.

a) Physiologie

Der menschliche Stimmapparat wird mit einem Blasinstrument verglichen. Der Luftstrom tritt aus dem "Windraum", gebildet von Lunge, Bronchien und Trachea, durch einen Spalt, der von schwingungsfähigen Lippen, den Stimmbändern, begrenzt ist, aus. Dabei versetzt er die Luftsäule der angrenzenden Lufträume wie Pharynx, Mundhöhle und Nase in Schwingungen. Die Schwingung der Stimmbänder bestimmt die Frequenz und damit die Höhe des Tones. Die Form, Größe und Anordnung der angrenzenden Räume bestimmt die Obertöne und Teiltöne des Klanges und somit den Charakter und die Klangfarbe der Stimme. Der Kehlkopf ist ein sehr vielseitiger Schallerzeuger, da viele Komponenten in weiten Grenzen variabel sind. Veränderlich sind der Anblasedruck im Windraum, die Spannung der Stimmlippen, die Form der Stimmritze, ihre Lage zum Ansatzrohr, die Form und Größe des Ansatzrohres. Eine große Rolle spielt bei all diesen Faktoren die Vermittlung des Nervensystems zur raschen Muskelwirkung. Es liegt eine reflektorisch gesteuerte Regulation mit weitgehender Einschaltung des Gehörs vor. Die besonderen Leistungen des menschlichen Kehlkopfes sind nicht in seinem Bau, sondern in der höheren Differenzierung des Zentralnervensystems begründet.

Die Öffnung der Stimmritze beginnt gleichzeitig vorn und hinten, zuletzt trennt sich die Mitte. Bei explosionsartigem Austritt der Luft aus der Glottis sinkt der Druck darüber ab, und Luft strömt von unten nach. Sie saugt die medialen Flächen der Stimmlippen wieder zusammen, wobei die Stimmlippenmasse im Moment des Rückschwungs diese Bewegung unterstützt. Die Flächen legen sich von unten nach oben aneinander.

Über den primären "Kehlkopftton" liegen verschiedene Anschauungen vor. BECKMANN und OEKEN finden anhand von elektroakustischen Analysen von Vokalen, daß er ein undifferenziertes Klanggemisch mit Frequenzen bis zu 1000 Hz darstellt. ARNDT hingegen meint, daß die primären Kehlklänge periodische Klanganteile bis 8000 Hz und höher haben.

Wirbelbildungen und Selbsterregung von Schwingungen dicht oberhalb der Stimmritze sind wichtig für die Klangfarbe und den Stimmcharakter. Der Luftstrom hat hier die größte Geschwindigkeit. Auch von der Epiglottis ist der Stimmklang mitabhängig. Ist sie gesenkt, enthält die Stimme mehr

hohe und höchste Teiltöne, ein Aufrichten verschleiert die Stimme.

Der Einsatz der Stimme kann auf drei verschiedene Weisen erfolgen. Beim harten Stimmeinsatz ist die Stimmritze schon vor der Phonation geschlossen. Ein plötzlicher, harter Stimmbeginn durch Glottissprengung setzt ein.

Beim weichen Stimmeinsatz beginnen die Stimmlippen zu schwingen, bevor sie die endgültige Stimmstellung erreicht haben, er geht also von der offenen Phase aus.

Wenn vor Erreichen der Stimmstellung bereits die Expirationsluft durch die Glottis strömt, spricht man von gehauchtem Stimmeinsatz. (v. ESSEN)

Die Anblaseluft und die Stimmlippen stellen ein ungedämpftes System dar, das mit der Luft des oberen Ansatzrohres so schwach gekoppelt ist, daß selbst im Resonanzfall keine wesentliche Frequenzstörung zustande kommt. Die Verstärkung der Frequenzen bis zu 300 Hz erfolgt im Windraum, die der höheren im Ansatzrohr.

Nach v. ESSEN hat jeder Resonanzraum einen gewissen Frequenzbereich, in dem er noch oder schon anspricht auf die äußeren Schwingungen. Dieser ist abhängig von der Gestalt, der Eigentonhöhe, dem Material, der Konsistenz der Wandungen - je weicher sie sind, desto weiter ist der Frequenzbereich und umso schneller klingt der Eigenton nach Aufhören der Erregung ab. Mund und Nase sind solche stark dämpfenden Resonatoren mit großer Resonanzbreite.

b) Physikalisches

Im ungedämpften System kann die Resonanzkurve bis Unendlich anwachsen. Das Maß für die Dämpfung ist das Amplitudenverhalten zweier aufeinanderfolgender Schwingungen. Der Logarithmus naturalis dieses Verhaltens wird als Dekrement (d) bezeichnet. Für die Mund- und Rachenhöhle beträgt d $0,1 - 0,4$, was einer flachen Resonanzkurve entspricht.

Die Verbindung benachbarter Schwingungssysteme, wie sie Kehlkopf und Ansatzrohr mit den verschiedenen Höhlen des Kopfes darstellen, wird als Koppelung bezeichnet. Beim Übergang der Luft vom Kehlkopfgenerator in das

gekoppelte Mund-Rachenhöhlensystem findet eine Veränderung des Klangspektrums statt. Wenn an zwei Koppelfrequenzen die Teiltöne verstärkt sind und die übrigen gedämpft werden, erhält man die charakteristischen Vokalspektren. Die willkürliche Veränderung der Koppelfrequenzen ergibt die verschiedenen Vokale. Die Frequenzen der Amplitudenmaxima bestimmen die Vokalfarbe. Diese sind die Formanten. Ihre Lage ist für die einzelnen Vokale erforscht. "U", "o" und "a" haben jeweils einen Formanten; "u" von 200 - 400 Hz, "o" von 400 - 600 Hz und "a" von 800 - 1200 Hz. "E" hat zwei Formanten, der erste reicht von 400 - 600 Hz, der zweite von 2200 - 2600 Hz. Ebenso ist es bei "i", dessen unterer Formant von 200 - 400 Hz reicht und der obere von 2800 - 3200 Hz. (TRENDLENBURG)

Als Klang wird die Superposition von Sinusschwingungen bezeichnet, deren Frequenzen in einem ganzzahligen Verhältnis stehen. Es sind die harmonischen Klänge der Natur- oder Obertonreihe. Ein Geräusch stellt eine Mischung von harmonischen und unharmonischen Teiltönen dar, wie es bei den Zischlauten der Fall ist.

Die Schallstärke wird aus der Geschwindigkeit des Hin- und Herschwingens der einzelnen Luftteilchen und der Dichteänderung der Luft, also der Schalldruckänderung, hergeleitet.

Schallintensität ist die Schalleistung bezogen auf den Flächeninhalt bzw. die Energie, die pro Sekunde durch einen cm^2 strömt. Sie ist dem Quadrat des Schalldrucks proportional und wird in Watt/cm^2 gemessen. Das Verhältnis von zwei Schallenergien wird durch den Lautstärkenunterschied als $10 \log I_1/I_2$ mit dem Maß Dezibel (dB) ausgedrückt. Das absolute Maß für die Lautstärke ist das Phon, bezogen auf einen Normalton von 1000 Hz. Die logarithmische Skala reicht von 0 Phon ($2 \cdot 10^{-4}$ mbar = Hörschwelle) bis zur Schmerzschwelle von 130 Phon. Die normale Singstimme hat 65 - 85 Phon. Der Eichpunkt der Skala liegt bei 74 Phon, was einem mbar entspricht. Die Lautheit wird in Sonen gemessen, wobei die festgelegte Einteilung 40 Phon = 1 Sone lautet. (LUCHSINGER). Die maximale Lautheit der Stimme ist von der Tätigkeit der Stimmlippen und der Verstärkung im Ansatzrohr abhängig.

Der tiefste Ton der menschlichen Stimme liegt bei 43 Hz, der höchste bei 12607 Hz, was einem Stimmumfang von etwa 6 Oktaven entspricht. (FLETCHER)

Die Tonhöhe ist von der Dauer der einzelnen Schwingungsperioden abhängig, die Intensität von der Größe der Schwingungsamplitude, und die Klangfarbe von der Form der Schwingung. (v. ESSEN)

IV. Möglichkeiten zur Verstellung der Sprechstimme

1) Physiologische Verstellmöglichkeiten

Abgesehen von den krankhaften Veränderungen der Stimme gibt es eine ganze Reihe von Möglichkeiten, die Stimme im Rahmen des Physiologischen willkürlich zu verändern. Da die Stimmfarbe durch Größe, Gestalt und Struktur des Stimmorgans und des Ansatzrohres bedingt ist, liegen hier die Ansatzpunkte für die Stimmverstellung. Ein hohler, dumpfer, gaumiger Klang entsteht z.B. durch Hochziehen des weichen Gaumens, was automatisch eine Tieferstellung des Kehlkopfes zur Folge hat und ein gespanntes Herunterdrücken der Zunge. Einen näselsnden Stimmklang kann man durch ein tiefstehendes Gaumensegel, hochgezogenen Kehlkopf, Verengung des Schlundeinganges und Hochlage des Zungenkörpers erreichen. Beim nasalen Tonfall ist bei Normalstellung von Ansatzrohr, Kehlkopf und Schlundeingangsweite der weiche Gaumen gesenkt. (v. ESSEN)

Die Verstellung der Stimmhöhe ist durch Veränderung der Stimmritzenöffnung und der Spannung der Stimmbänder bis zu bestimmten, von Person zu Person unterschiedlichen Grenzen zu erreichen. Die Artikulation und somit die Form des vorderen Ansatzrohres kann in verhältnismäßig weiten Grenzen variiert werden. Es kann sehr schlecht, schlecht, verwaschen, mäßig, durchschnittlich, gut, scharf, sehr scharf oder forciert artikuliert werden. Dazu gehören auch Verstellung der Mundart und Annahme eines fremdländischen Akzents. Sprechmelodie und Rhythmus können im Sinne einer gesamtphonischen Verregelung abgeändert werden, was eine rational gesteuerte Bremsung und Stereotypisierung in der Lautabfolge bedeutet. Gut modifizierbar ist die Lautstärke, weniger die Größe des Stimmvolumens. (R. FÄHRMANN). Eine besonders gute Stimmverstellung gelingt durch die sogenannte Ösophagussprache, bei der verschluckte Luft durch den Ösophagus zurückgepreßt und artikuliert wird.

2) Künstliche Verstellungsmöglichkeiten

Unter künstlicher Stimmverstellung versteht man eine Verstellung, die nur durch Manipulation von außen erreichbar ist, sei es durch Gegenstände oder sonstige Maßnahmen.

Beispielsweise wird durch einen Druck, der von außen auf die Stimmlippen ausgeübt wird, ein gespreßter, verspannter, heiserer Stimmklang erzeugt, die sogenannte Taschenbandstimme. Die Taschenbänder übernehmen hierbei die Aufgabe der außer Funktion gesetzten Stimmbänder, was die Veränderung zur Folge hat. Diese Methode wird vorwiegend von Menschen bevorzugt, die zu chronischer Heiserkeit und Stimmbelagtheit neigen.

Sehr häufig wird Verstellung der Stimme und Sprache durch Vorhalten der Hand oder eines Tuches beim Sprechen praktiziert. Dadurch kommt eine Dämpfung des Stimmklanges mit dumpfem, verwaschenem Klangbild zustande. Primär ist aber der gesamte Artikulationseffekt betroffen. Es treten auch Veränderungen des Timbre und der gesamten Ausdrucksbasis ein, besonders wenn zusätzlich eine weitere Verstelltechnik angewandt wird. Insgesamt können höchstens vier bis fünf phonetische Qualitäten gleichzeitig modifiziert werden. Unter dem Einfluß von Alkohol erfährt die Stimme ebenfalls eine Veränderung. Sie wird entspannter, lauter, tiefer, voller, in der Färbung gaumiger und nasaler, was auf phonischen Lösungserscheinungen beruht. Bei Volltrunkenheit allerdings kann es zu extremer Stimmstellung kommen, die durch krasse Klangverzerrung, übermäßige Dehnung und unartikulierte Lautfolge gekennzeichnet ist. Als weitere Möglichkeit zur Verstellung wird die Wirkung von Fremdkörpern in der Mundhöhle erwähnt. Ein Stift zwischen den Zahnreihen, ein Stein oder Speise im Mund verändern beim Sprechen die Artikulation zum Undeutlichen, Verwaschenen; auf die Stimme üben diese Mittel keinen Einfluß aus.
(R. FÄHRMANN und H. HABERSBRUNNER)

V. Verfahren zur Identifizierung von Sprechern aus der Literatur

In Amerika hat sich die "visible speech"-Methode im Dienst der kriminaltechnischen Forschung bereits bewährt.

L.G. KERSTA berichtet über Versuchsreihen mit "Stimmabdrücken", die er 1962 mit der sonographischen "contour-display"-Methode durchführt. Er läßt z.B. von fünf Sprechern zehn Schlüsselworte jeweils viermal über Telefon sagen; zuerst einzeln, dann innerhalb von zwei fünf-Wort-Sätzen und noch einmal während eines Textes, in dem jedes Schlüsselwort zweimal erscheint. Die Sonagramme werden von sieben Testpersonen, die auf sonographische Identifizierung geschult worden waren, nach dem subjektiven optischen Eindruck verglichen und zugeordnet. Von diesen Prüfern machen vier keinen Fehler, zwei machen einen Fehler und einer zwei. In weiteren solcher Serien erhält KERSTA Richtigkeitswerte von 97 % und 98 % bei 5000 Entscheidungen. Er kann sogar ermitteln, daß Stimmverstellungen geschulter Schauspieler und Bauchredner nicht so gut sein können, als daß sie mit Hilfe der Sonographie nicht zu identifizieren wären. Auch mechanische Mittel zur Stimmveränderung machen ein Sonagramm nicht unkenntlich. H. HABERSBRUNNER unternimmt 1968 einen Versuch zur Feststellung der Brauchbarkeit von Stimmanalysen. Er verwendet nicht den Sonographen, sondern ein Kombinationsgerät mit Tonfrequenz-Analysator der Firma Rhode u. Schwarz. Die Analysen werden vom Oszillographenschirm abgelichtet und die Photos ausgewertet. Die Versuche gelten vornehmlich der Funktion der Apparatur, es werden aber auch Verstellungen mit vorgehaltenem Taschentuch, Speisen im Mund und einem Bleistift zwischen den Zähnen aufgenommen, wobei man sich auf die Worte "ja" und "bitte" beschränkt. Eine Diskussion dieser Ergebnisse wird nicht durchgeführt.

Auf Wunsch der kriminaltechnischen Untersuchungsstelle in Hamburg macht O.v.ESSSEN 1966 gezielte Aufnahmen zur Feststellung von Sprecheridentität mit dem Sonographen in der Streifendiagrammtechnik. Er verwendet ausschließlich das Wort "Apotheke", das er von fünf Sprechern an zwei verschiedenen Tagen je einmal aufnimmt. Er läßt das Testwort jedesmal mit natürlicher, unverstellter Stimme und Sprechweise direkt auf die Magnettonscheibe des Sonographen sprechen. Die subjektiv-optische Auswertung durch zehn völlig ungeübte Testpersonen ergibt eine Wahrscheinlichkeit der richtigen Zuordnung von 97,5 %. Zur weiteren Untersuchung der Partialfrequenzverteilung stellt er zunächst zeichnerisch die Übereinstimmungen von Resonanzstreifen bei allen Vokalen für Ausspruch I und II fest. Diese setzt er ins Verhältnis zu den maximal möglichen Übereinstimmungen bei jedem Vokalpaar. Die so erhaltenen Wahrscheinlichkeitswerte für die Übereinstimmung der Frequenzbänder vergleicht er unter allen Vokalen und

stellt fest, daß statistisch kein wesentlicher Unterschied zwischen den Vokalen besteht. Sie können als einheitlich in Bezug auf ihre Übereinstimmungsquote zwischen Sprechprobe I und II angesehen werden. Der Vergleich der Wahrscheinlichkeitsquotienten jedes Vokals bei den verschiedenen Sprechern untereinander ergibt, daß im Mittel von 5,5 möglichen Formantübereinstimmungen nur 1,27 - 1,43 gemeinsam vorgefunden werden. Das besagt, daß die vokalischen Sprechklänge bei verschiedenen Sprechern verschieden sind. Die Gestaltung der Laut- und Pausendauer scheidet er als Kriterium aus, da ihre Untersuchung nicht ergiebig ist. Diese Versuchsreihe berücksichtigt keine willkürlichen Stimm- und Artikulationsänderungen, sondern verweist bezüglich der Anzahl der Sprecher, der Variation von Text und Stimme und der Art der Aufnahmen auf künftige Untersuchungen.

C EIGENE UNTERSUCHUNGEN

I. Material und Versuchsbeschreibung

1) Die sonographischen Aufnahmen

Für die eigene Untersuchungsreihe werden hauptsächlich sonographische Aufnahmen gemacht. Dazu benutzen wir zwei verschiedene Geräte.

a) Der Sonagraph

Der Hauptteil der Diagramme wird in der Stimm- und Sprachabteilung der Universitäts-HNO-Klinik München unter Verwendung des Sona-Graph Model Recorder 662 B der amerikanischen Firma "Kay Electric Company" hergestellt. Der vorhandene Einschub liefert nur lineare Streifen-sonagramme. Die Apparatur analysiert außerdem die Gesamtamplitude und im "Sectionner" jede beliebige Intensitäts-Frequenzbeziehung. Letzteres wird nicht angewendet.

In der kriminaltechnischen Abteilung des Bayerischen Landeskriminalamtes können mit einem anderen Sonagraphen der "Kay Electric Company", dem Sona-Graph 6061-B und dem zugehörigen Einschub, der "Contour Display Unit" 6070-A, einige weitere Aufnahmen im Relief-Verfahren gemacht werden.

b) Die Aufnahmetechnik

Das Testwort wird jeweils zweimal über Mikrophon direkt auf die Magnettonscheibe des Sonographen gesprochen, und zwar kurz hintereinander, um innerhalb der 2,4 sec. der Papierlänge zu bleiben. Jedes Signalpaar wird unmittelbar nach der Aufnahme analysiert, jeweils in den 45- und 300 Hz Bandfilterbreiten mit totaler Spektrumaufzeichnung von 85 - 8000 Hz und zusätzlich mit einer Dehnung des unteren Frequenzbereiches auf das Doppelte, also von 85 - 4000 Hz. Die Einstellung des Intensitätsniveaus schwankt je nach Lautstärke der Sprechprobe zwischen 5,5 und 9,5 (Mark Level). Für jeden Sprecher wird auch die Gesamtamplitude aufgezeichnet, die mit einer Dämpfung von 10 dB erscheint.

Im Landeskriminalamt werden die Aufnahmen ganz genauso durchgeführt; es kann von jedem Sprecher nur ein Reliefdiagramm angefertigt werden. Die Intensitätseinstellung beträgt durchschnittlich 2,5 (Mark Level).

2) Die Aufnahmen mit dem Kathodenstrahloszillographen

In der Klinik werden von einigen der Sprecher ergänzende Aufnahmen mit dem Kathodenstrahloszillographen und zugehörigem Formantfilterverstärker gemacht.

a) Der Kathodenstrahloszillograph

Wir benutzen das in der Stimm- und Sprachabteilung der Universitäts-HNO-Klinik zur Verfügung stehende Zehnkanal-Großbildoszilloskop Typ 43 BN 900 der Firma Knott Elektronik, München. Von der gleichen Firma stammt das nach E. LOEBELL speziell für diesen Zweck konstruierte Zusatzgerät für die Frequenzanalyse. Dieser Formantfilterverstärker zerlegt die Signalfrequenzen in acht Bereiche: von 50-300, 300-600, 600-900, 900-1250, 1250-1650, 1650-2100, 2100-2600, 2600-10000 Hz. Da die oberen Frequenzbereiche nicht so bedeutend für die Auswertung sind, kann das letzte Intervall so groß gewählt werden. An die einzelnen Bereiche schließt sich eine Darstellung des Gesamtspektrums an. Die Fotoaufnahmen werden mit einer "cinerollex"-Kamera der Fa. Linhof, Format 56 x 72 mm gemacht. Der Kodak-Film Tri X Pan ist 70 mm breit und hat eine Empfindlichkeit von 27 Din.

b) Die Aufnahmetechnik

Das Testwort wird jeweils einmal über Mikrofon in 30 cm Abstand vom Mund gesprochen. Nach Einstellung der Ablaufgeschwindigkeit erfolgt mit Beginn der Sprechprobe Aufschlüsselung, Aufzeichnung und Ablichtung automatisch.

II. Angaben über die Versuchspersonen

1) Anzahl

In der Klinik werden insgesamt 81 Sprecher sonagraphisch getestet, bei 12 von ihnen machen wir außerdem Aufnahmen mit Oszillographen. Im Landeskriminalamt können von 25 Versuchspersonen sonagraphische Aufnahmen gemacht werden.

2) Alter

Bei den Gewährspersonen handelt es sich ausschließlich um Erwachsene im Alter zwischen 15 und 62 Jahren. Das Durchschnittsalter beträgt etwa 40 Jahre.

3) Geschlecht

Die Gesamtzahl von 106 sonagraphischen Aufnahmen gliedert sich in 30 weibliche und 76 männliche Sprecher auf. Bei den Oszillographenaufnahmen beträgt die Anzahl der weiblichen Sprecher 3 und die der männlichen 9.

4) Sprachzugehörigkeit

Bis auf zwei Sprecher sind alle Testpersonen deutschsprachig. Die beiden Ausländer, ein Tscheche und ein Ungar, haben in der deutschen Aussprache einen leichten Akzent. Etwa 20 der Versuchspersonen sprechen deutlich bayerische Mundart. Exakte hochdeutsche Aussprache liegt bei 15 Testpersonen vor. Alle übrigen artikulieren ohne auffällige Dialektfärbung in durchschnittlicher Weise.

5) Sprachgeschulte Testpersonen

Von allen getesteten Personen sind 16 durch ihren Beruf auf dem Gebiet der Stimme und Sprache als sprachgeschult anzusehen. Die Sonagramme dieser Sprecher fallen weit deutlicher, klarer begrenzt und ausgeprägter aus als die der meisten ungeschulten Sprecher. Vor allem in dem zweiten Versuch mit verstellter Stimme erweisen sich die Sonagramme der sprachgeschulten Sprecher als recht unterschiedlich gegenüber der ersten natürlichen Sprechprobe, was bei den meisten der übrigen ungeschulten Sprecher nicht der Fall ist.

6) Ungeschulte Sprecher

Alle übrigen Sonagramme stammen von Sprechern ohne spezielle sprach- und stimmliche Ausbildung. Keiner von ihnen weist Sprachfehler oder stimmliche Defekte auf. Außer Stimm- und Sprachgestörten wird in zufälliger Auswahl jede Person getestet, die sich zu der Sprechprobe bereit erklärt. Alle ungeschulten Sprecher haben noch nie einen Sonagraphen besprochen, kurze Übungszeit und einige Wiederholungen sind in fast jedem Fall notwendig. Da aus zeitlichen Gründen keine längeren Übungszeiten möglich sind, und sich die Verstellung der Stimme unmittelbar an die natürliche Aussprache anschließen soll, ergeben sich bei der Stimmverstellung oft Schwierigkeiten und ungenügende Abänderungen. Letzteres hat auch teilweise im Unvermögen der Sprecher seinen Grund. In den Sonagrammen zeigen sich diese Schwächen deutlich, die Analysen weisen entsprechend wenig Unterschiede auf.

III. Wortwahl und Programm

In Fortführung der Versuchsreihe OTTOS v. ESSEN über "Identifizierung von Sprechstimmen nach sonographischen Registrierungen" und einiger Analysen von E. LOEBELL zu diesem Thema benutzen wir für unsere Untersuchungen ebenfalls das Wort "Apotheke". Ein Einzelwort hat den Vorteil der besseren Lesbarkeit des Sonagramms. Die einzelnen Laute und ihre Formanten sind leicht aufzufinden und sehr gut vergleichbar, zumal sie bei diesem kurzen Wort auch bei zweimaliger Aussprache auf einem Registrierpapier festgehalten werden können. Das Wort "Apotheke" besitzt viele verschiedene, durch Explosivlaute gut voneinander abgegrenzte Vokale. Sie erschei-

nen im Sonagramm besonders deutlich und eignen sich gut zu einer vermessenden Auswertung. "Apotheke" läßt sich leicht artikulieren und fördert durch die harte Konsonantenskandierung eine klare Aussprache heraus.

Wie im vorherigen Abschnitt bereits erwähnt, soll das Wort "Apotheke" von jedem Sprecher zunächst in natürlicher Stimmlage und Sprechweise gesprochen werden. Unmittelbar daran anschließend ist eine möglichst weitgehende Verstellung von Stimme oder Artikulation anzuwenden oder eine gute Verstellung durch mechanische Mittel zu erreichen. Die Sprechproben werden in den beschriebenen Weisen aufgenommen.

IV. Methodik

1) Veränderung der Tonhöhe

Die Sprecher werden aufgefordert, zur zweiten Sprechprobe ihre Stimme möglichst weit nach oben oder unten, je nach besserem Vermögen, zu verstellen. Dabei ergibt sich keine eindeutige Verstellungstendenz bei den verschiedenen Geschlechtern. Jeweils 50 % der männlichen und der weiblichen Sprecher verstellen ihre Stimme nach oben, die anderen 50 % nach unten.

2) Tarnung

Verschiedene Sprecher bevorzugen statt der Tonhöhenveränderung beim verstellten Sprechen eine Dämpfung der Stimme mit Hilfe der Hand oder eines Taschentuches. Dabei tritt generell eine Abschwächung der oberen Resonanzen ein. Wenige bedienen sich auch einer Verstellung durch näselnde Aussprache. Bei der Auswertung werden beide Verstellungsarten mit zu den veränderten Tonhöhen gezählt.

3) Fremdkörperwirkung

Zur Vervollständigung der Versuche werden mit 9 Versuchspersonen weitere Aufnahmen gemacht, um die Fremdkörperwirkung im Mund beim Sprechen zu untersuchen. Die Testpersonen nehmen einen Stift zwischen die Zähne und artikulieren so das Wort "Apotheke".

4) Veränderung der Mundart

Dieselben 9 Sprecher verstellen bei einem zweiten Versuch die Mundart willkürlich und vollständig. Nachdem der erste Ausspruch bei allen hochdeutsch vorliegt, verändern sie jetzt nach bayerisch, sächsisch und plattdeutsch.

Die beiden letzten Versuche werden gesondert ausgewertet.

V. Auswertung

1) Fragestellung

Von der vorliegenden Literatur angeregt, interessiert die Frage, ob es mit Hilfe der Sonographie möglich ist, anhand der Stimmanalysen manifeste Erscheinungsbilder in der menschlichen Stimmaufzeichnung festzustellen und so die Sonographie zu einem Mittel der Personenidentifizierung zu machen. Speziell interessiert hier die Verteilung, Form und Lagebeziehung der Resonanzen und Formanten der Vokale und ihre eventuelle Änderung bei einem willkürlichen Verstellen der Stimme, das möglichst bis zur Unkenntlichkeit für das Gehör reichen sollte. Die Untersuchung wird sich im wesentlichen auf diesen Bereich beschränken. Die Konsonanten des Wortes "Apotheke" werden wegen der schlechten metrischen Erfassbarkeit ihrer sonographischen Darstellung nicht berücksichtigt.

2) Auswertung der gewonnenen Analysen

a) Subjektiver visueller Bildvergleich für Streifen- und Reliefdiagramme

Aus dem gesamten Untersuchungsmaterial der Streifen- und Reliefdiagramme werden jeweils 5 Blätter zufällig ausgesucht, die natürliche Sprechweise von der verstellten getrennt und zur Erkennung mit Handzeichen versehen. Die jeweiligen 5 unverstellten Sprechproben werden der Reihe nach ausgelegt, die jeweils zugehörigen 5 verstellten ungeordnet daneben. Die Aufgabe für unvoreingenommene Versuchspersonen besteht darin, die verstellten Aussprüche auf Grund der Ähnlichkeit des Klangbildes den entsprechenden unverstellten zuzuordnen. Diesen Test führen für die Streifensonagramme 11 und für die Reliefdiagramme 10 Personen durch. Für 5 richtig zuge-

ordnete Blätter gibt es die Höchstzahl von 4 Richtigkeitspunkten.

a1) Ergebnisse

Folgende Ergebnisse sind zu verzeichnen:

Streifendiagramm:	Punktzahl	0	1	2	3	4	Durchschnittliche Punktzahl: 2,64
	Personenzahl	0	3	0	6	2	
=====							
Reliefdiagramm:	Punktzahl	0	1	2	3	4	Durchschnittliche Punktzahl: 2,3
	Personenzahl	0	3	3	2	2	

a2) Statistischer Schluß

Für die Streifendiagramme beträgt die Varianz der Werte 1,25. Die Wahrscheinlichkeit der richtigen Zuordnung von Bild I zu Bild II liegt bei 66 %. Die Berechnung der Normalvariablen "K" und der dazugehörige Tabellenwert gibt in % die Fälle einer großen Zahl von Testpersonen an, deren Punktzahl noch außerhalb des vom Mittel am weitest entfernten Einzelwertes liegen wird, die in diesem Fall also weniger als einen Punkt erreichen würden. Das sind hier 7,35 %. Die Varianz der Reliefdiagrammwerte beträgt 1,22. Hier liegt die Wahrscheinlichkeit richtiger Zuordnung nach dem vorliegenden Ergebnis nur bei 57,5 %. Das besagt nach der Berechnung von "K" für eine große Zahl von Fällen, daß mindestens 14,7 % noch unter einem Punkt liegen werden.

b) Vergleichende Vermessung der Resonanzstreifen mit statistischer Prüfung für

b1) gleiche Sprecher mit verschiedener Sprechprobe

Mit Hilfe einer Schablone, auf der vertikal die Frequenz in KHz (0,005-2,0) und horizontal die Zeit in sec. (0-2,4) angegeben sind, werden die Resonanzbereiche ausgemessen und in entsprechender Lage auf Millimeterpapier übertragen, so daß jeweils die Vokale der ersten und der zweiten Sprechprobe von "Apotheke" nebeneinander liegen. Anschließend werden durch Querverbindungen die Übereinstimmung von Resonanzstreifen festgestellt, wobei angrenzende Resonanzen wegen der Ungenauigkeit der Schablone als nicht

Übereinstimmend betrachtet werden. Nach Auszählung der Übereinstimmungen und der maximalen Resonanzbandzahl, gleichgültig ob sie im ersten oder im zweiten Ausspruch vorkommt, wird für jeden Vokal der Wahrscheinlichkeitswert nach der Definition von LAPLACE gebildet:

$$P(A) = \frac{g}{m} = \frac{\text{Anzahl der Fälle, da A eintritt}}{\text{insgesamt mögliche Fälle}}$$

So erhält jedes Vokalpaar jedes Sprechers seinen Quotienten.

Als Mittelwerte (M) und Varianzen (s^2) ergeben sich für die Vokale von "Apotheke":

Für A:	Für O:	Für E ₁ :	Für E ₂ :
$M = 0,69$	$M = 0,64$	$M = 0,68$	$M = 0,67$
$s^2 = 0,044$	$s^2 = 0,042$	$s^2 = 0,054$	$s^2 = 0,042$

Der Unterschied der Mittel ist nach dem Test der Mittelwerte mehrerer Normalverteilungen zwischen den einzelnen Vokalen unerheblich, was bedeutet, daß die Übereinstimmungsquoten bei den verschiedenen Vokalen im wesentlichen gleich sind.

Nach Aufstellen einer Häufigkeitstabelle erhält man die Anzahl der Sprecher auf die Zahl der Übereinstimmungen in % für jeden Vokal. Da die Vokale als Gesamtheit betrachtet werden können, werden sie in der folgenden Tabelle zusammengefaßt, und es wird die durchschnittliche Anzahl der Sprecher ermittelt, die auf die in 10 %-Gruppen zusammengefaßten Übereinstimmungen fallen.

Mittelwert der Personenzahl	Übereinstimmungen aller Resonanzen
0	0 - 20 %
2,25	20 - 30 %
5,5	30 - 40 %
7,0	40 - 50 %
14,0	50 - 60 %
12,75	60 - 70 %
11,0	70 - 80 %
13,5	80 - 90 %
1,25	90 - 100 %
11,75	100 - .. %

Unter der Annahme, daß zu einer Identifizierung der Sprecher ein Höchstsatz von 40 % Nichtübereinstimmungen noch toleriert werden kann, lauten

die weiteren Ergebnisse:

bei 28,75 von insgesamt 79 Sprechern stimmen 0 - 60 % der Resonanzstreifen überein, bei 50,25 von insgesamt 79 Sprechern stimmen sie zu 60 - 100 % überein; in Prozenten ausgedrückt zeigen 36,5 % der Personen 0 - 60 % Übereinstimmung und 63,5 % 60 - 100 %. Das heißt, daß die Mehrzahl der Fälle auf Grund der Meßergebnisse zu identifizieren ist.

b2) verschiedene Sprecher mit unverstellter Sprechprobe

Die nächste Untersuchung der übertragenen Frequenzbänder gilt dem Vergleich des ersten, also unverstellten Ausspruchs aller Sprecher untereinander. Es ist festzustellen, ob und inwieweit ein Unterschied in der Formant- und Resonanzverteilung bei verschiedenen Personen gegenüber der Stimmverstellung einer Person besteht. Auch hier werden Übereinstimmungen und maximale Möglichkeiten ausgezählt und für jeden Vokal die Wahrscheinlichkeitswerte gebildet. Aus der Masse der Werte werden wegen der Notwendigkeit der Gruppengrößengleichheit zufällige 78 Werte für jeden Vokal ausgewählt. Zunächst wird der Test der Mittelwerte mehrerer Normalverteilungen durchgeführt. Es ergeben sich die Mittelwerte und Varianzen:

Für A:	Für O:	Für E ₁ :	Für E ₂ :
M = 0,61	M = 0,54	M = 0,59 *	M = 0,58
s ² = 0,035	s ² = 0,043	s ² = 0,039	s ² = 0,054

Der Test ergibt keinen wesentlichen Unterschied zwischen den einzelnen Mittelwerten, was wiederum eine ziemlich gleiche Übereinstimmungsquote der verschiedenen Vokale bedeutet.

Ob der Unterschied der Mittel zwischen der zweimaligen Sprechweise von "Apotheke" einer Versuchsperson und den unverstellten Sprechproben zweier verschiedener Versuchspersonen statistisch wesentlich ist oder nicht, wird mit dem Test zum Vergleich der Mittelwerte zweier Normalverteilungen geprüft und für jeden Vokal gesondert durchgeführt. Das Ergebnis fällt für jeden Vokal positiv aus, was bedeutet, daß die Hypothese $M_1 = M_2$ zugunsten der Annahme, daß M_1 wesentlich größer als M_2 ist, verworfen werden muß. Das besagt wiederum, daß die Wahrscheinlichkeit, daß ein Sprecher bei Stimmverstellung wesentlich mehr Übereinstimmungen der Resonanzbänder zeigt als zwei verschiedene Individuen beim Ausspruch des gleichen Wortes aufweisen, genügend groß ist.

Die Auszählung sämtlicher gewonnenen Werte erfolgt in Häufigkeitstabellen für jeden Vokal einzeln. Es ergeben sich folgende Werte:

Vokal	Gesamtzahl d. Fälle	Zahl d. Fälle m. 0-60 % Übereinstimmung	in %	Zahl d. Fälle m. 60-100 % Übereinstimmung	in %
A	1672	847	50,6	825	49,4
O	1782	1050	58,9	732	41,1
E ₁	1799	1003	55,7	796	44,3
E ₂	1724	985	57,2	739	57,2

Eine Vergleichstabelle der Prozentzahlen von Versuch I (dem Vergleich bei-der Sprechproben eines Sprechers) und Versuch II (dem Vergleich der natür-lichen Sprechweise zweier verschiedener Sprecher) soll die Unterschiedlich-keit der Werte verdeutlichen.

	0-60 % Übereinstimmungen		60-100 % Übereinstimmungen	
	Versuch I	Versuch II	Versuch I	Versuch II
A	23,0 %	50,6 %	77,0 %	49,4 %
O	44,5 %	58,9 %	55,5 %	41,1 %
E ₁	44,5 %	55,7 %	55,5 %	44,3 %
E ₂	35,4 %	57,2 %	64,6 %	42,8 %

c) Vergleichende Vermessung der durch Stift zwischen den Zähnen oder durch Mundart verstellten Sprechweisen von "Apotheke"

Bei diesen Versuchen werden die Resonanzstreifen ebenfalls vermessen und mit denen der normalen Sprechprobe verglichen. Übereinstimmungen werden in der graphischen Übertragung wieder mit Verbindungslinien gekennzeichnet und ausgezählt. Auch sie werden zu den maximalen Übereinstimmungsmöglichkeiten ins Verhältnis gesetzt. Für die Aussprache mit Stift werden folgende Durch-schnittsquotienten ermittelt:

M "A": 0,6, M "O": 0,59, M "E₁": 0,65, M "E₂": 0,56

Das Mittel der Mittelwerte beträgt 0,6, was einer durchschnittlichen Übereinstimmungsrate von 60 % entspricht. Für diese Versuche werden keine statistischen Berechnungen durchgeführt. Bei den vorliegenden Werten kann für die 9 Aufnahmen mit Artikulationsbeeinträchtigung durch Stifteinwirkung eine Identifizierbarkeit angenommen werden.

Etwas abweichend liegen die Meßwerte für die verstellte Mundart:

M "A": 0,45 M "O": 0,67 M "E₁": 0,56 M "E₂": 0,62

Die Verstellung durch Änderung der Mundart gelingt in sonographischer Hinsicht besser. Hier liegt die Zahl der Personen, die von 0 - 50 % übereinstimmen, um 3 höher als bei Verstellung mit Stift zwischen den Zähnen. Unter der Annahme, daß die Vokale gleich zu bewerten sind, ist diese Verstellungsart mit einem Durchschnittsquotienten von 0,57 nicht zu identifizieren.

d) Amplituden- und Lautdauergestaltung

Die Aufzeichnung der Gesamtamplitude mit einer generellen Dämpfung von 10 dB im oberen Drittel des Registrierpapiers wird für die Mehrzahl der Sprecher durchgeführt. Sie stellt die undifferenzierte Gesamtintensität für jeden Laut dar. Auf eine Auswertung wird verzichtet, weil diese Amplitude zur Frage der Resonanzverteilung nichts beiträgt und als Kriterium an sich zu geringe Aussagekraft besitzt.

Auch auf zeitliche Dauer der Laute und Pausen wird, im Gegensatz zu v. ESSEN, nicht eingegangen, da die Gestaltung der Lautdauer durch die stark begrenzte Zeit von 2,4 Sekunden sehr vorgeschrieben ist und auf keinen Fall den individuellen Besonderheiten entsprechen kann. Hinzu kommt die durch die notwendige Reaktionsschnelligkeit bedingte psychische Belastung der Versuchspersonen, die lediglich die Konzentration auf die geforderte Verstellung erlaubt. Von unserer Seite werden die Versuchspersonen auf diese Möglichkeit auch nicht aufmerksam gemacht. Gelegentliche Änderungen des Sprachrhythmus im Vergleich zur natürlichen Sprechweise dürften rein zufällig und für das Ergebnis weniger bedeutend sein.

e) Vergleichende Vermessung der Reliefdiagramme

e1) Frequenzmessung mit dem Schreibstift des Sonagraphen und statistischer Schluß

Für die Auswertung der Reliefdiagramme stellt sich das Problem der metrischen Erfassung von Intensitätsfeldern. Zunächst wird versucht, die Felder größter Intensität, die den Formanten entsprechen, in ihrer Frequenzausdehnung zu messen. Dazu steht ein zusätzlicher Einschub des Sonagraphen zur automatischen Frequenzzählung bei Ablauf des Sonagramms zur Verfügung. Leider war das Gerät bei der Anfertigung der Sonagramme noch nicht vorhanden, so daß nun die Notwendigkeit besteht, die Frequenzmessung nachzuvollziehen. Das Wiedereinspannen des Sonagramms, das Ausrichten auf die Null-Linie und das Abgreifen der Feldgrenzen mit dem handgeführten Glühstift bergen viele Fehlerquellen, so daß die auf drei Kommastellen genaue Frequenzanzeige hier ziemlich unnötig ist.

Die relativen Maßwerte dürften aber noch genau genug sein. Auch hier wird der Quotient aus tatsächlichen Übereinstimmungen und den maximal möglichen gebildet. Die Mittelwerte und Varianzen lauten:

Für A:	Für O:	Für E ₁ :	Für E ₂ :
M = 0,64	M = 0,57	M = 0,69	M = 0,80
s ² = 0,133	s ² = 0,115	s ² = 0,107	s ² = 0,063

Der F-Test der Mittelwerte besagt trotz des abweichenden Mittelwertes von E₂, daß kein wesentlicher Unterschied zwischen den Werten für die einzelnen Vokalpaare besteht. Die vorliegenden Werte stimmen zu durchschnittlich 67,5 % überein, eine Identifizierung der Sprecher auf Grund der Verteilung der stärksten Intensitäten ist damit möglich.

Eine reine Frequenzmessung ist für Reliefdiagramme aber nicht zweckmäßig, da durch die besondere Art der Aufzeichnung in ihrer Lage ähnlich verteilte Felder geschaffen werden, die durch ihre flächige Ausdehnung sehr viele gemeinsame Punkte aufweisen. Viel unterschiedlicher gestalten sich hier die Flächenformen der verschiedenen Intensitätsgrade, die die wesentlichen Merkmale der verschiedenen Sprecher bilden.

e2) Planimetrische Messungen

Einen Versuch zur objektiven Darstellung der Flächenverhältnisse stellt die planimetrische Auswertung dar, wobei die Flächengrößen der stärksten und der darauffolgenden schwächeren Intensitätsstufen jeweils zueinander ins Verhältnis gesetzt werden, und nach übereinstimmenden Quotienten gesucht wird. Zuerst benutzen wir ein Planimeter der Firma Ott/Kempton, Typ 10, das aber wegen seiner groben Skaleneinteilung bei der Kleinheit mancher Flächen ungeeignet ist. Die Messung mit einer selbstgefertigten Schablone in mm²-Einteilung erweist sich als günstiger. Das Ergebnis bleibt aber enttäuschen. Von jeweils 20 ausgewerteten Fällen stimmen für "A" 7 Fälle oder 35 % überein, für "O" 4 Fälle oder 20 %, für "E₁" 7 Fälle oder 35 % und für "E₂" 7 Fälle oder 35 %. Eine Übereinstimmungsrate von durchschnittlich 31,2 % reicht für eine Identifizierung in keiner Weise aus. Eine wirklich günstige Auswertung dieser Diagramme kann wohl nur eine Sonograph-Computer-Kombination liefern.

f) Vergleich der Oszillographenbilder

f1) Messung der Amplitudenmaxima

Die Photographien der oszillographischen Analysenergebnisse legen die natürliche und die verstellte Sprechprobe auf getrennten Aufnahmen vor. Die Ausschläge in den Frequenzbereichen stellen die Partialintensitäten der Vokallaute dar. Der jeweils größte Ausschlag der einzelnen Lautdarstellungen wird mit Hilfe eines Stechzirkels gemessen und der Wert von einer Millimeterskala abgelesen. Die Kanäle 1, 7 und 8 fallen zur Beurteilung weg, da in ihnen keine oder nur ganz geringe Ausschläge zu registrieren sind. Die gemessenen Werte erstrecken sich also über einen Frequenzbereich von 300 - 2100 Hz. Die in ihrer größten Amplitude übereinstimmenden Werte der einzelnen Vokale von Ausspruch I und Ausspruch II werden ausgezählt. Dabei sind $\pm 0,2$ mm Differenz noch als übereinstimmend anzunehmen. Von den insgesamt 14 ausgewerteten Analysen ergibt sich folgendes Bild:

Anzahl der Übereinstimmungen						
Frequenz (Hz)	Kanal	A	O	E ₁	E ₂	gesamt
300 - 600	2	8	6	3	6	23
600 - 900	3	2	4	2	3	11
900 - 1250	4	4	7	5	6	22
1250 - 1650	5	3	4	2	5	14
1650 - 2100	6	7	8	3	5	23

Die höchste Zahl von Übereinstimmungen weist "0" auf, die niedrigste "E₁". Betrachtet man alle Vokale als Einheit und zählt sämtliche Übereinstimmungen pro Person aus, so ergibt sich:

Zahl d. Übereinstimmungen	1	3	5	6	8	9	10	13
Zahl d.-Personen	1	3	2	2	2	2	1	1
Quote	0,05	0,15	0,25	0,3	0,4	0,45	0,5	0,65

Da für jeden Vokal insgesamt 20 Übereinstimmungen möglich gewesen wären, erhält man die untenstehenden Wahrscheinlichkeitsquotienten. Ihr Mittelwert beträgt 0,34, was einer durchschnittlichen Übereinstimmungsrate von 34 % entspricht. Damit ist auf diese Weise eine Identifizierung der oszillographischen Bilder nicht möglich.

D DISKUSSION

I. Diskussion der Ergebnisse

Sowohl der für Streifen- als auch für Reliefdiagramme durchgeführte visuelle subjektive Bildvergleich mit durchschnittlichen Punktzahlen von 2,54 und 2,30 von 4,0 möglichen ergibt eine schlechte Identifizierbarkeit von Sonagrammen für ungeübte Testpersonen. Die Wahrscheinlichkeit der richtigen Zuordnung liegt für diese Werte bei 60 % und 57,5 % im Mittel.

Im Gegensatz dazu erhält OTTO v. ESSEN in seiner "Apotheken"-Versuchsreihe bei gleichen Testbedingungen durchschnittlich 3,9 Richtigkeitspunkte und damit eine Wahrscheinlichkeit der richtigen Zuordnung von 97,5 %. Allerdings stammen seine Sonagramme von 5 Sprechern mit jeweils zwei natürlichen, ungestellten Sprechproben, die lediglich an verschiedenen Tagen aufgenommen worden sind. Dafür ist die Wahrscheinlichkeit der richtigen Zuordnung also sehr hoch. Das besagt schon, daß durch Stimmverstellung doch recht merkbare Veränderungen im Klangbild stattfinden können. Ein geeignetes Kriterium für eine wissenschaftliche Auswertung ist so eine Prüfung jedoch nicht, da die Fähigkeit, Sonagramme auf Grund subjektiver visueller Erfassung richtig zuzuordnen, durch fortgesetzte Übung und Sorgfalt beim Vergleich sehr erhöht werden kann, andererseits aber auch die Gefahr der Fehleinschätzung groß ist.

Auch L.G. KERSTA erzielt in seinen Versuchsreihen mit dem subjektiv-visuellen Bildvergleich sehr positive Ergebnisse. Wie bereits ausgeführt, schult er seine Testpersonen im Lesen von Sonagrammen, so daß bei Einzelworten, Sätzen und Texten mit 97 %, 98 % und 99 % richtig getroffenen Entscheidungen ein hervorragendes Ergebnis erzielt wird.

Die Vermessung der Partialfrequenzen und anschließende statistische Prüfung ergibt bei OTTO v. ESSEN mit 99 % Wahrscheinlichkeit, daß die Mittelwerte der einzelnen Vokalquotienten eines Sprechers im wesentlichen übereinstimmen, die verschiedenen Vokale des Wortes "Apotheke" in ihren Übereinstimmungsraten also als gleichwertig zu betrachten sind.

Zu diesem Ergebnis sind wir bei gleicher Auswertmethode sowohl für Streifen- als auch Reliefdiagramme ebenfalls gekommen. Die zusätzliche, nicht statistische Auswertung gilt der Ermittlung der Einzelwerte der Übereinstimmungen. Für Streifensonagramme ergibt die Überwiegende Mehrheit der Sprecher (63,5 % gegenüber 36,5 %) Übereinstimmung der Frequenzbänder zwischen 50 % und 100 %. Das käme einer gewissen Identifizierung gleich, wenn 60 % Übereinstimmung noch als genügend zu einer Erkennung betrachtet werden. Da zu einer sicheren juristischen Identifizierung jedoch 99 % Übereinstimmung gegeben sein müssen, kann eine Identität mit an Sicherheit grenzender Wahrscheinlichkeit aus der Überwiegenden Zahl der Fälle nicht abgeleitet werden. Andererseits muß aber ab 60 % Übereinstimmung bereits mit einer Sprecheridentität gerechnet werden, was bei Verdacht durch andere Indizierung zu stützen ist. Als alleiniger Schuldbeweis ist die Sonographie aber in den vorliegenden Fällen unzulänglich. Bei den Reliefdiagrammen stimmt ebenfalls die Mehrheit (60 % gegenüber 40 %) der Sprecher zu 60 % und darüber in den Feldern stärkster Intensität überein.

Den Vergleich verschiedener Sprecher untereinander führt O. v. ESSEN so durch, daß er die Gemeinsamkeiten der bei zweimaligem Sprechen übereinstimmenden Gebiete erhöhter Resonanz bei seinen 5 Gewährspersonen miteinander vergleicht. Er findet von durchschnittlich 5,5 Übereinstimmungen nur 1,27 bis 1,43 im Mittel gemeinsam vor. Das besagt, daß die vokalischen Sprechklänge verschiedener Personen verschieden sind. Wegen der Unterschiedlichkeit der Aussprüche eines Sprechers ziehen wir nicht die Übereinstimmungen zum Vergleich heran, sondern vergleichen jeweils die erste, unverstellte Sprechprobe jedes Sprechers miteinander. Der Test

der Mittelwerte ergibt, wie bei O. v. ESSEN auch, wieder keinen wesentlichen Unterschied der Vokalmittel. Der für jedes Vokalpaar einzeln durchgeführte Test der Mittelwerte zweier Normalverteilungen ergibt bei jedem Vokal eine statistisch wesentlich größere Übereinstimmungsrate für Sprecher mit Stimmverstellung als für zwei verschiedene Personen. Die Auszählung der Einzelwerte zeigt, daß die Mehrheit der Sprecher hier Übereinstimmungen in den Frequenzbändern von 0 - 60 % aufweist (55,6 % gegenüber 44,4 %). Damit ist eine Identifizierung nach unserer Definition nicht möglich. Das Ergebnis stimmt mit dem von JOOS, POTTER, PETERSON, BARNAY und O. v. ESSEN überein. Verschiedene Personen produzieren also immer unterscheidbare Klangbilder. Die Untersuchung kann ebenfalls nachweisen, daß bei ungenügender Verstellung von Stimme oder Sprechweise bei einem Individuum sichere Erkennungsmerkmale im Klangbild stets vorhanden sind. Ob bei akustisch unkenntlicher Stimmverstellung durch sprachgeschulte oder verstellungsbegabte Testpersonen die Analyse immer in genügendem Maß von den Normwerten abweicht, ist wegen der geringen Anzahl der dazu vorliegenden Fälle nicht nachweisbar. Die Anzeichen deuten aber darauf hin, daß ein Erkennen mit visuell-metrischen Mitteln nicht mehr sicher möglich ist.

Ein Verundeutlichen der Artikulation durch Fremdkörperwirkung weist im Sonagramm keine größeren Unterschiede auf, während eine völlige Änderung der Artikulation, wie sie bei guter mundartlicher Verstellung durch Veränderung der Klangform, Klangfülle, Betonung, Rhythmik und Verteilung der Resonanzanteile mit entsprechenden Änderungen im Ansatzrohr verbunden ist, eine erheblichere Abweichung von den Werten der Norm bedingt. Besonders die Vokale werden von solchen Abwandlungen betroffen; in unserem Versuch weist der Vokal "A" die größten Abweichungen auf.

Die Intensitätsfeldmessungen, die für Reliefdiagramme auf Anregung von H. HABERSBRUNNER durchgeführt werden, weist so starke Abweichungen zwischen den beiden Aussprüchen des Wortes "Apotheke" auf, daß diese Methode für eine Auswertung nicht geeignet erscheint. Die beste Auswertmethode dafür ist, abgesehen von Computerauswertung, der visuelle Vergleich durch geübte Personen, der durch Messungen gestützt wird.

Auch der bei der oszillographischen Aufnahmetechnik ermittelte durchschnittliche Mittelwert von 0,34 stellt eine zu geringe Zahl von Übereinstimmungen dar, als daß eine Identifizierung möglich wäre. Die Ausehläge sind teilweise sehr schlecht zu erkennen, was die Messung behindert oder verfälscht.

Trotz der Aufgliederung in die acht Filterbereiche bleibt die Analyse zu komplex und ist für eine genauere Auswertung unzugänglich. Für Stimmidentifizierung muß diese Methode ausscheiden.

II. Ergebnis für die Fragestellung

Anhand der hier durchgeführten Versuchsauswertung muß festgestellt werden, daß für jedes Individuum typische Kennzeichen im Stimmanalysenbild existieren. Aus unserer Versuchsreihe kann nicht hervorgehen, daß diese spezifische Lage und Form der Formanten und Nebenformanten durch Stimmverstellung grundlegend geändert werden kann. Eine Identifizierung wird in vielen Fällen von Stimmverstellung auch tatsächlich möglich sein, besonders, wenn es sich um geeignete Stichproben aus zusammenhängendem Text handelt. Die Mehrzahl der Sprecher ohne besondere Stimm- und Sprachschulung ist nicht in der Lage, so vollkommen und anhaltend zu verstellen, daß sie in längeren Texten nie in ihren natürlichen Stimm- und Sprechmodus zurückfallen. Eine Annäherung an die Norm zeigt sich im Sonagramm deutlich in der Ähnlichkeit des Musters und der Frequenzverteilung. In diesen Fällen stellt die Sonographie ein Hilfsmittel zur Identifizierung von Sprechstimmen dar. Ausgesprochene Stimmverstellungen ergeben jedoch so veränderte Sonagramme, daß mit den herkömmlichen Methoden der Auswertung nicht mehr mit Sicherheit auf den Sprecher geschlossen werden kann.

E ZUSAMMENFASSUNG

Die in amerikanischem und deutschem Schrifttum aufgeworfene Frage nach der Identifizierbarkeit von Sprechstimmen mittels Sonographie veranlassen uns zu eigenen Untersuchungen.

In der Universitäts-HNO-Klinik und im Bayerischen Landeskriminalamt München werden insgesamt 106 erwachsene Sprecher getestet. Davon sind 76 Männer und 30 Frauen.

Wir wählen das Testwort "Apotheke". Jeder Sprecher wendet für Sprechprobe I des Wortes "Apotheke" seine unverstellte, natürlich Sprechweise an, für Probe II eine Verstellung.

Am häufigsten wird die Stimme nach "hoch" oder "tief" verstellt, einige tarnen die Stimme, wenige verstellen durch Näseln. 9 der Testpersonen bringen zusätzlich Verstellung mit einem Stift zwischen den Zähnen und Mundartabwandlung.

Jede Sprechprobe wird in das Mikrophon des Sonagraphen gesprochen und sofort analysiert. Es werden Streifensonagramme sowie Sonagramme in Relieftechnik hergestellt. 12 Sprecher wiederholen ihre Proben in einen Formantfilterverstärker, die Frequenzanalysen werden vom Oszillographenschirm abgeleitet.

Die Sonagramme werden auf verschiedene Weisen ausgewertet:

1. Der subjektive visuelle Vergleich von Sprechproben gleicher Sprecher, von jeweils 11 und 10 Testpersonen für je 5 Diagramme durchgeführt, ergibt die Wahrscheinlichkeit der richtigen Zuordnung für Streifensonagramme von 66 %, für Reliefdiagramme von 57,5 %.
2. Die metrisch-statistische Auswertung der Sprechproben gleicher Sprecher ergibt für die Mehrheit von 63,5 % der Sprecher mit Stimmverstellung und Stimmtarnung Übereinstimmung der Frequenzbänder zwischen 60 und 100 %. Die zugehörigen Oszillographenbilddauswertungen zeigen pro Personen im Mittel nur 34 % Übereinstimmung.

Die Verstellung mit Stift zwischen den Zähnen zeigt pro Person im Mittel 60 % Übereinstimmung, die Mundartänderung 57 %.

Die Reliefdiagramme ergeben für Stimmverstellung pro Person im Mittel 67,5 % Übereinstimmung.

3. Der metrisch-statistische Vergleich aller Sprechproben I der verschiedenen Sprecher untereinander ergibt einen statistisch signifikanten Unterschied zwischen den Vokalfrequenzlagen verschiedener Personen. 55,6 % der Sprecher zeigen Übereinstimmungen der Frequenzbänder zwischen 0 und 60 %.

Die Wahrscheinlichkeit, daß eine große Zahl von Sprechern in ihrer Stimmverstellung im Bereich zwischen 60 und 100 % mit der normalen Sprechweise übereinstimmt, ist in unserer Versuchsreihe statistisch genügend groß.

Die Wahrscheinlichkeit, daß eine große Zahl von verschiedenen Sprechern in ihrer normalen Sprechweise nur im Bereich von 0 - 60 % übereinstimmt,

ist hier statistisch ebenfalls genügend groß.

Die Ergebnisse stimmen im wesentlichen mit den in der Literatur vorliegenden überein.

F ANHANG

I. Außerliterarische Quellen

Neben der angegebenen Literatur wird die Kurzbeschreibung des "Sona-Graph Model Recorder", beigegeben von der Firma "Kay Electric Company", benutzt.

Die Kurzbeschreibung des Zehnkanaal-Großbild-Oszilloskops der Firma Knott wird ebenfalls verwendet.

Mündliche Mitteilungen erhalten wir von Herrn Dr. H. Habersbrunner über die Verwendung des Sonagraphen im zukünftigen Bankwesen;

von der cardiologischen Abteilung des Klinikums Rechts der Isar, München, über die Verwendung des Sonagraphen zu Herzschaallaufnahmen;

von einem Vertreter der Firma Nucleotron, München, über das Arbeitsprinzip des Sonagraphen;

von einem Vertreter der Firma Knott Elektronik, München, über die Bau- und Arbeitsweise des Formantfilterverstärkers, angeschlossen an das Großbildoszilloskop.

II. Abbildungen

Die folgenden Abbildungen stellen Beispiele für sonographische Aufnahmen in der Technik des Streifensonagramms und der des Reliefdiagramms dar. Erläuterungen folgen gesondert.

Abb. 1

Streifensonagramm von
"Apotheke", zweimal
von einem stimm- und
sprachlich ungeschul-
ten männlichen Spre-
cher in

1. natürlicher
 2. nach unten verstell-
ter Stimme
- gesprochen.

Deutlich sind die Über-
einstimmungen im Klang-
bild zu erkennen.

A p o t h e k e

2.

A p o t h e k e

1.

Abb.2

Streifenphonogramm

"Apotheke", zweimal

von einem stimm- und

sprachlich ungeschulten

männlichen Sprecher in

1. natürlicher

2. nach oben verster

ter Stimme

gesprochen.

Übereinstimmungen in

Klangstruktur sind

feststellbar.

TYPE B SONOGRAM ©

KA 2.

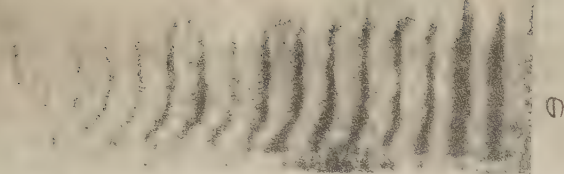
PINE BROOK, N. J.

9
k
e
th
o
p
H

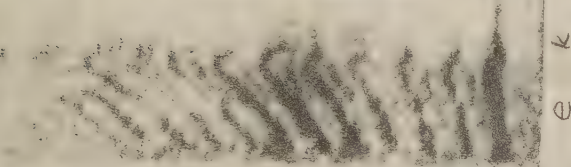
2.

9
k
e
th
o
p
P

1.



S



K



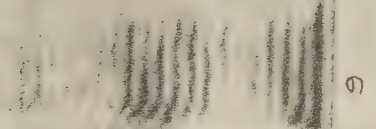
th



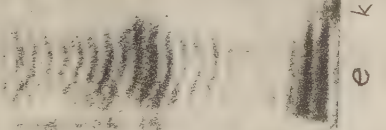
P

O

2.



S



K

e

th



P

O

1:

Abb.3

Streifensonagramm v

"Apotheke", zweimal

von einem stimm- und

sprachgeschulten we

lichen Sprecher in

1. natürlicher

2. nach oben verster

ter Stimme

gesprochen.

Hier weisen die Kla

bilder große Unter

schiede auf, eine I

tizifizierung ist nich

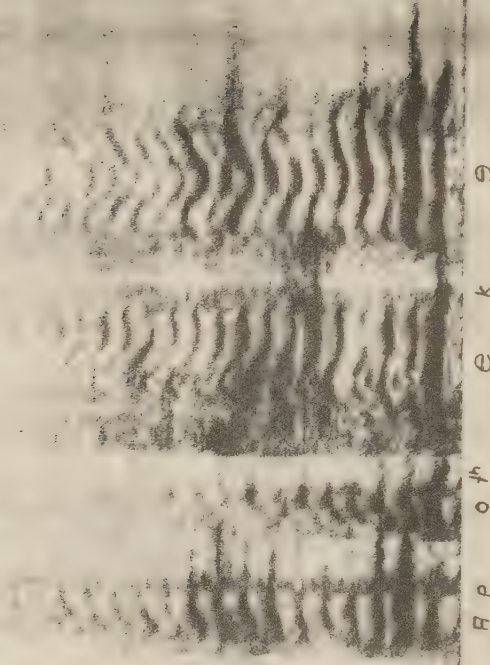
möglich.

Abb. 4

Streifensonogramm von
"Apotheke", zweimal
von einem stimm- und
sprachgeschulten männ-
lichen Sprecher in
1. natürlicher
2. nach oben verstell-

ter Stimme
gesprochen.

Auch hier ist das zwe-
te Klangbild stark ab-
gewandelt, es ist kei-
ne Identifizierung
möglich.



e

k

e

o

th

p

f

2.



e

k

e

o

th

p

f

1.

- 2.

4.

Form und Verteilung
der Intensitätsfelder
weisen Ähnlichkeiten
auf.

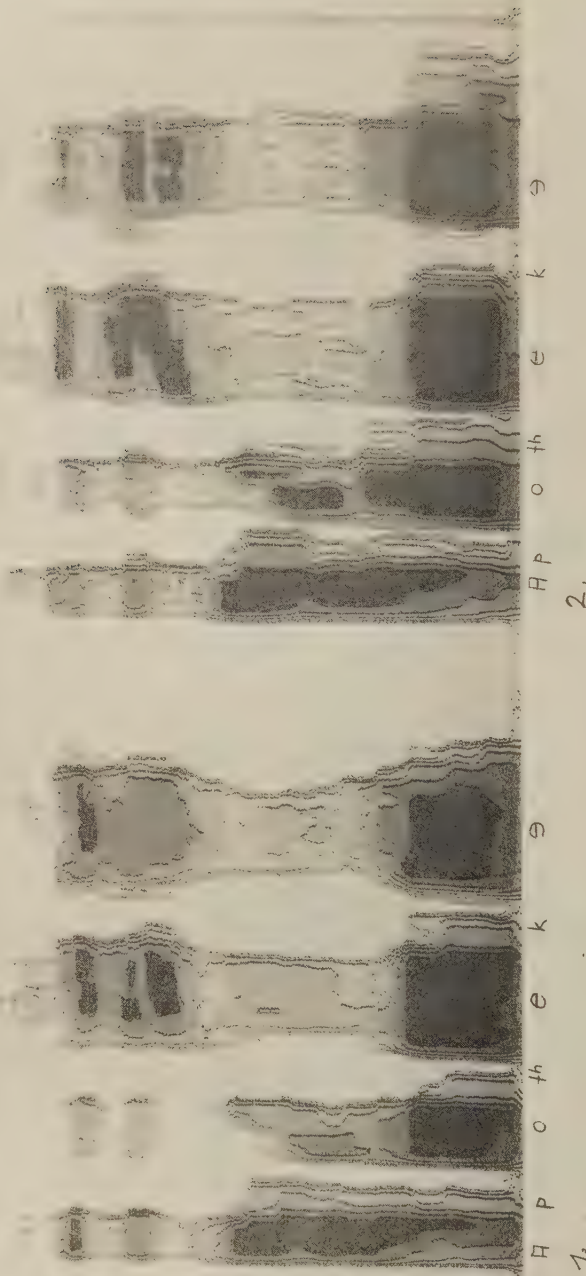


Abb.6

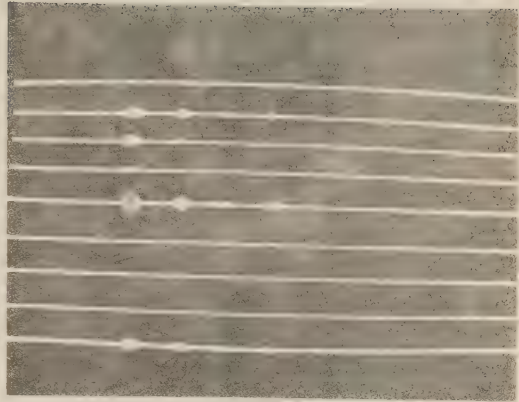
Reliefdiagramm von
 "Apotheke", zweimal
 von einem stimm- und
 sprachlich ungeschulten
 männlichen Sprecher in
 1. natürlicher
 2. nach oben verstärkter
 Stimme
 gesprochen.
 Die Übereinstimmung
 der Klangstrukturen
 sind hier sehr deutlich

Abb.7

Oszillographische Aufnahme von "Apotheke" nach Analyse im Formantfilterverstärker.

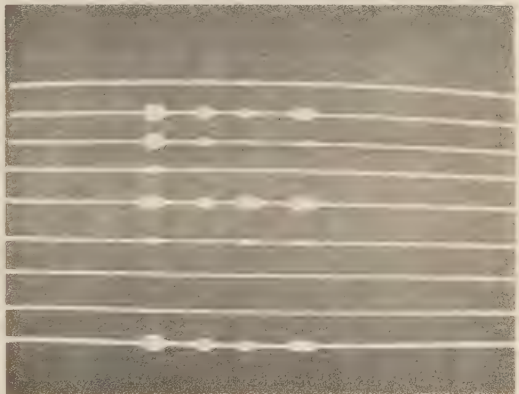
Unverstellte Stimme eines männlichen Sprechers.

Gesamtspektrum



Derselbe Sprecher mit nach unten verstellter Stimme. Die Amplitudenweiten weisen kaum Gemeinsamkeiten auf. Hier sind die Ausschläge allgemein höher als in Ausspruch 1.

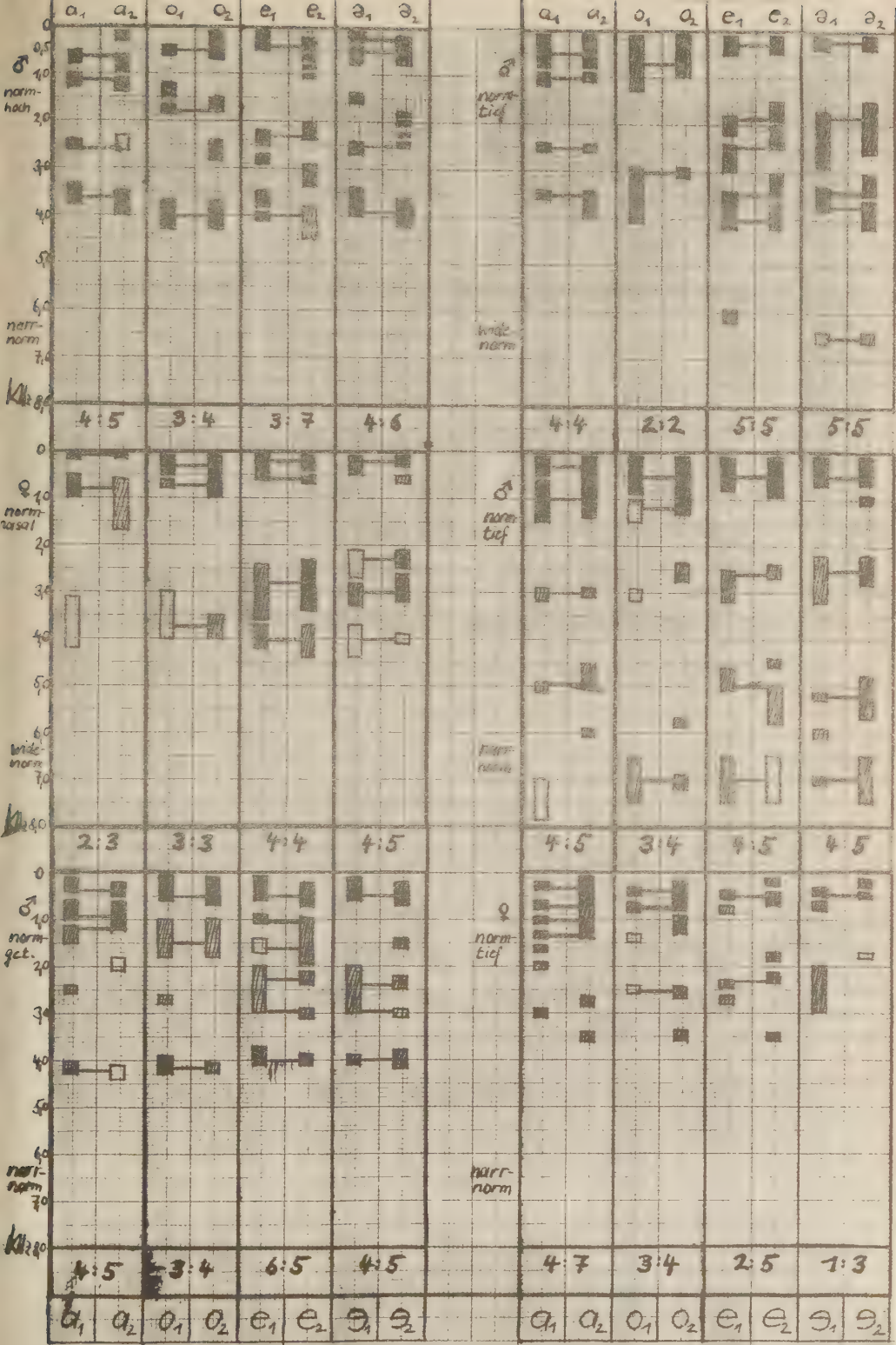
Gesamtspektrum

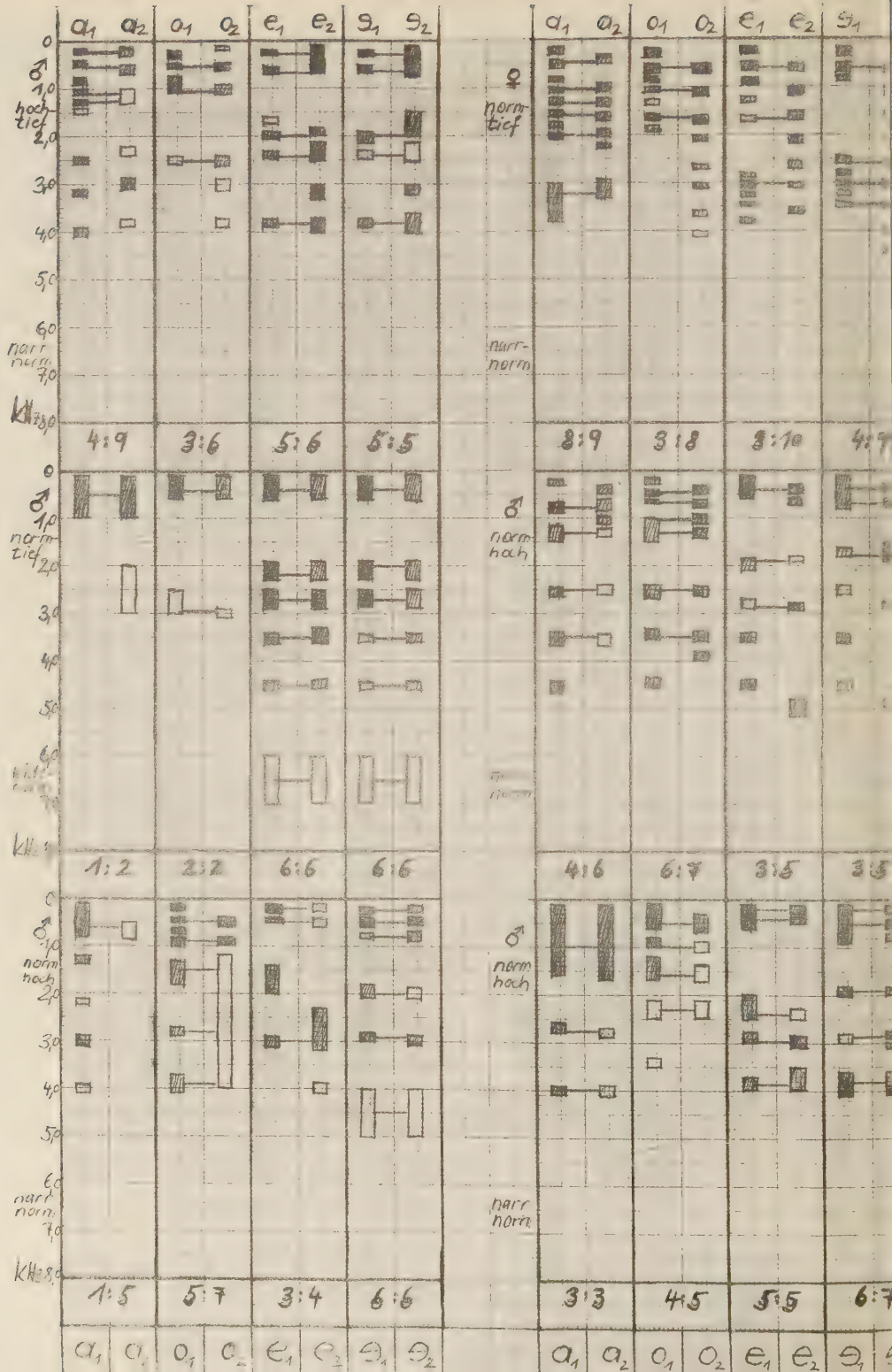


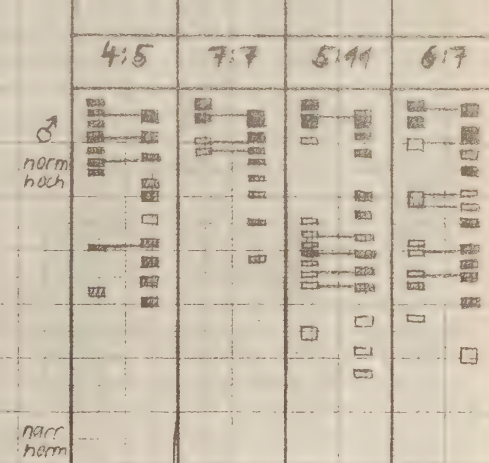
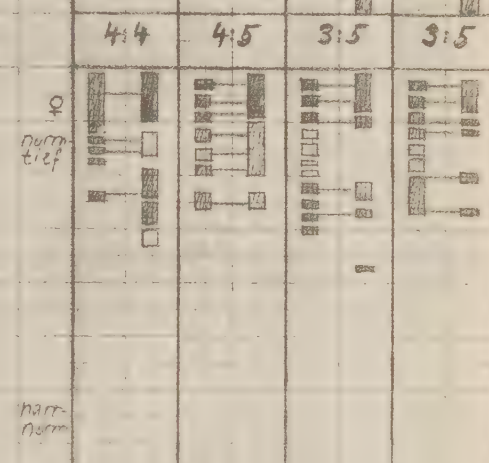
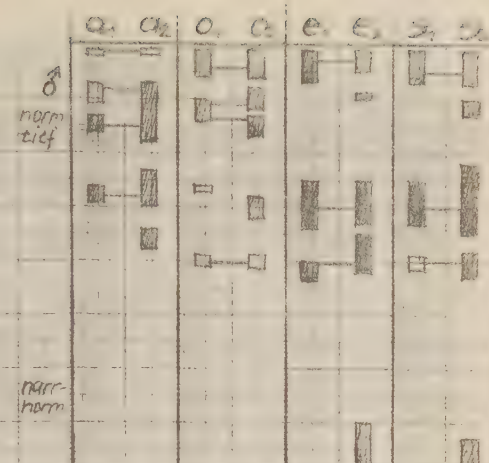
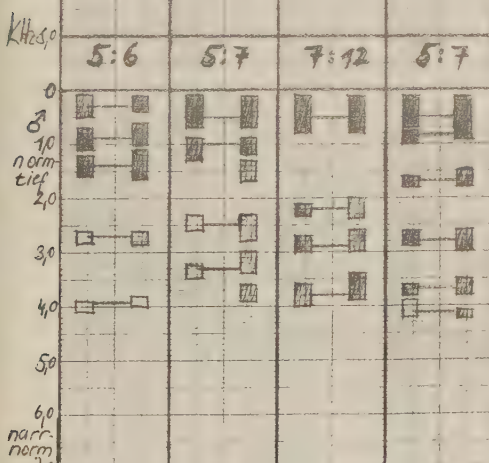
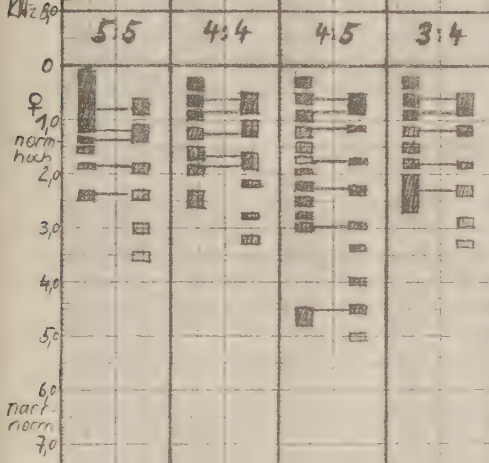
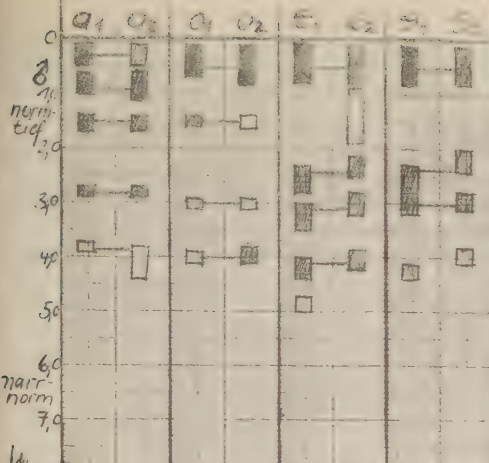
III. Graphische Darstellungen

Die Umzeichnungen der Sonagramme von "Apotheke" in natürlicher und verstellter Stimme oder Sprechweise zeigen die Partialfrequenzen der einzelnen Vokale. Sie sind entsprechend ihrer Frequenzlage im Sonagramm zwischen 85 und 8000 Hz angeordnet. Es liegen jeweils der gleiche Vokal der Sprechproben I und II jedes Sprechers nebeneinander. Übereinstimmungen der Resonanzbalken sind durch Querverbindungen gekennzeichnet. Ungefüllte Formen bedeuten eine schwache Ausprägung der entsprechenden Partialfrequenzen.

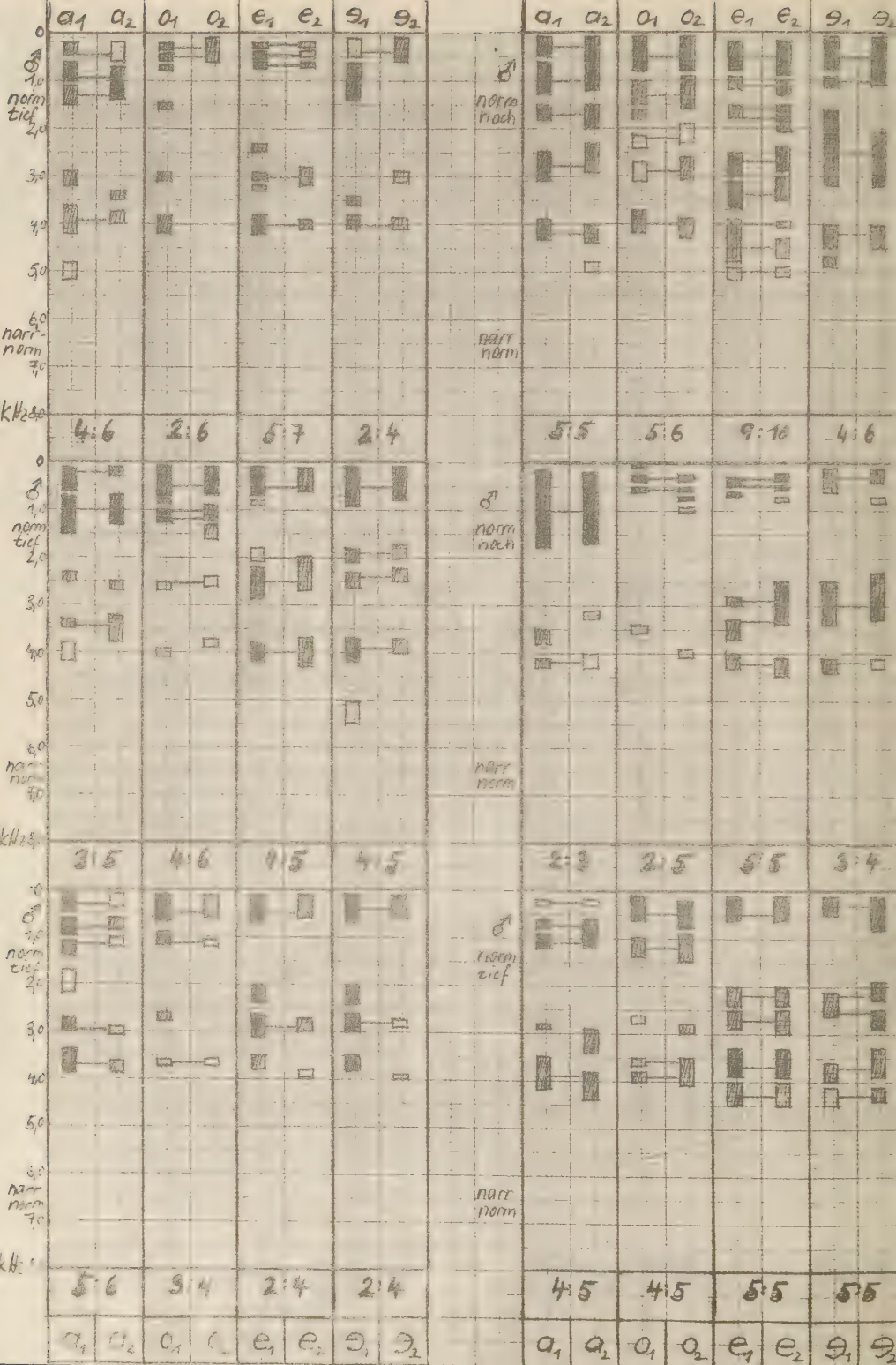
Das Verhältnis der tatsächlichen Übereinstimmungen zu den maximal möglichen stellen die Zahlen unterhalb der Vokalspalten dar.

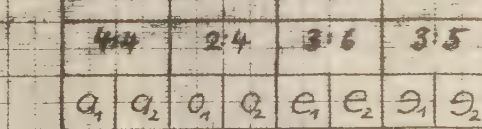
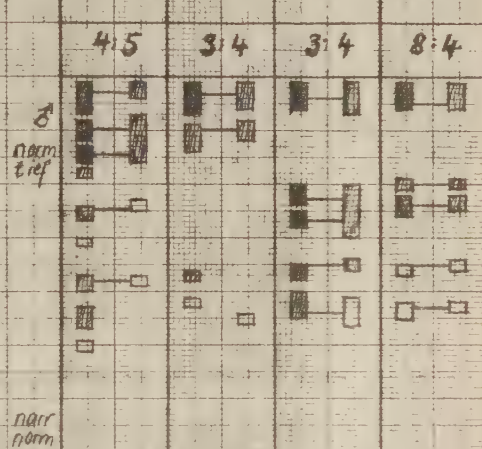
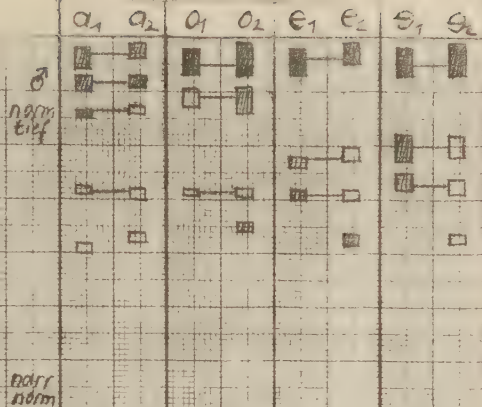
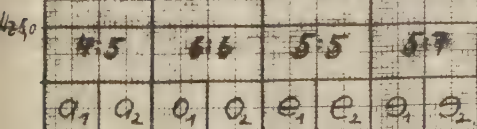


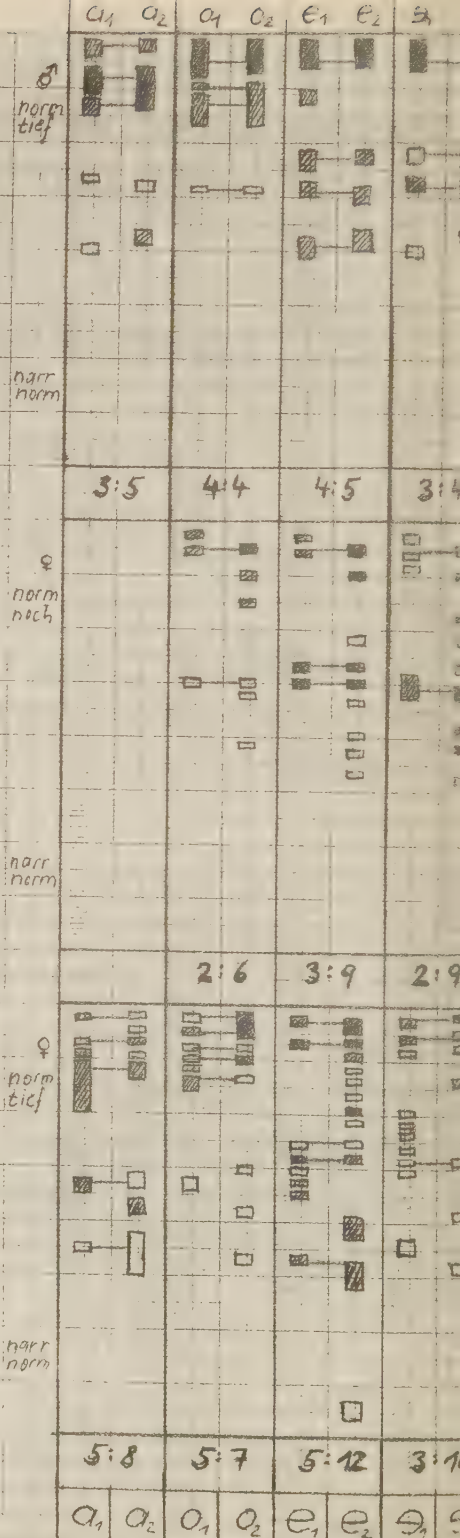
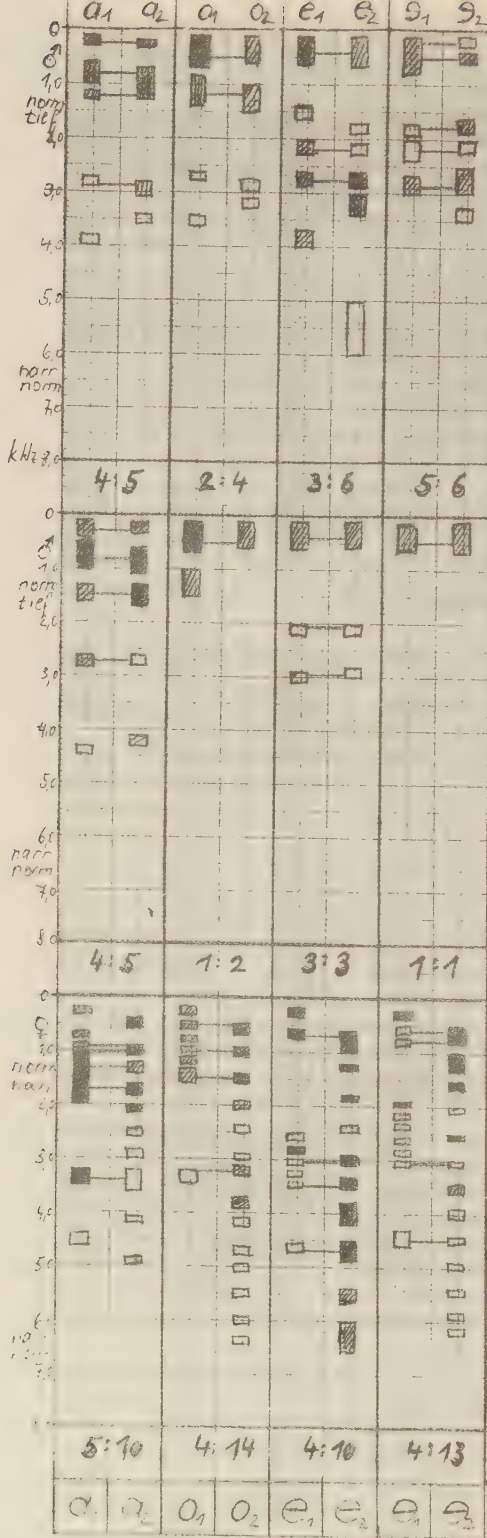




5:5	4:4	4:5	3:4	4:4	4:5	3:5	3:5
5:6	5:7	7:12	5:7	4:5	7:7	5:11	6:7
5:5	4:6	4:4	6:6	4:10	3:8	5:12	6:12
a ₁ a ₁	a ₁ a ₂	e ₁ e ₂	s ₁ s ₂	a ₁ a ₂	o ₁ o ₂	e ₁ e ₂	s ₁ s ₂

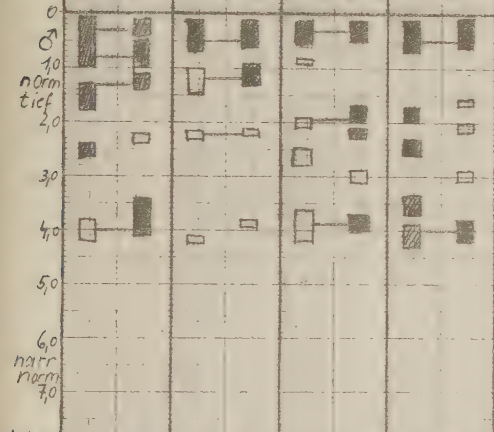




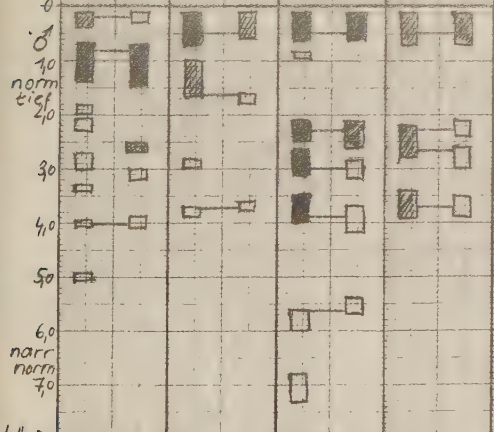




klz 8,0

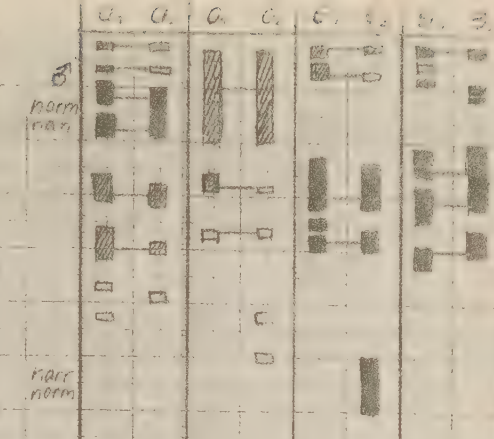


klz 8,0

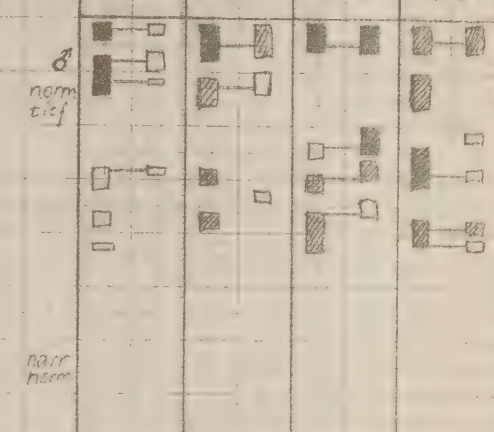


klz 8,0

Columns: a_1 , a_2 , e_1 , e_2



klz 8,0

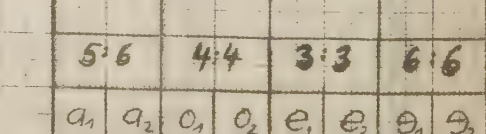
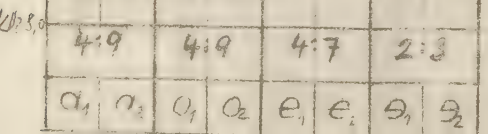
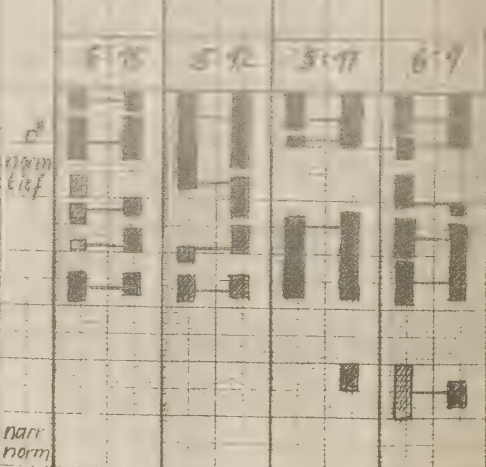
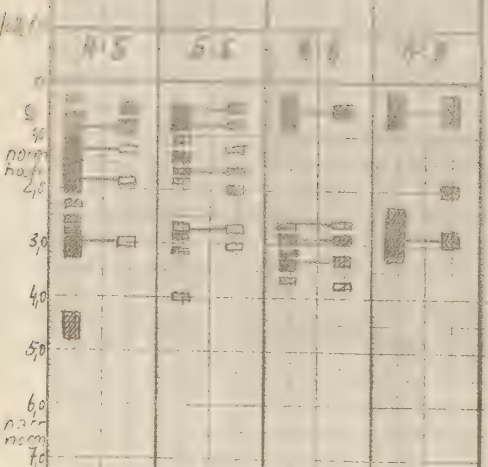
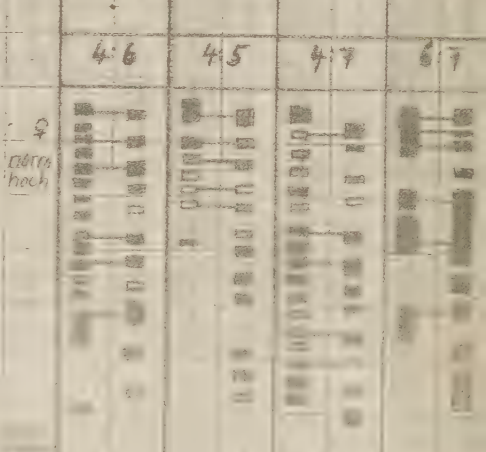
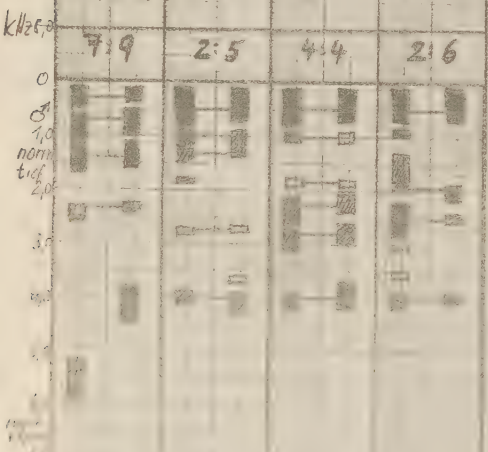
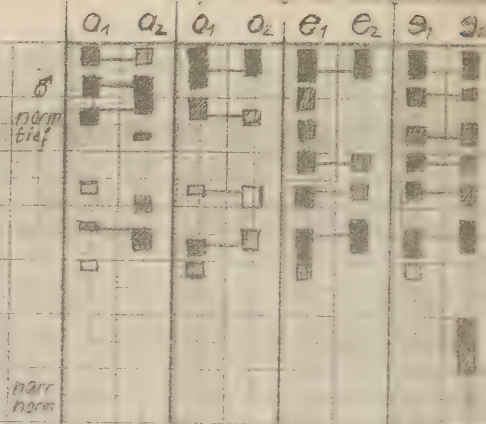
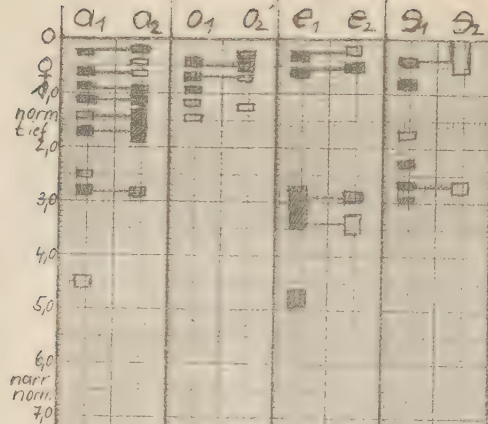


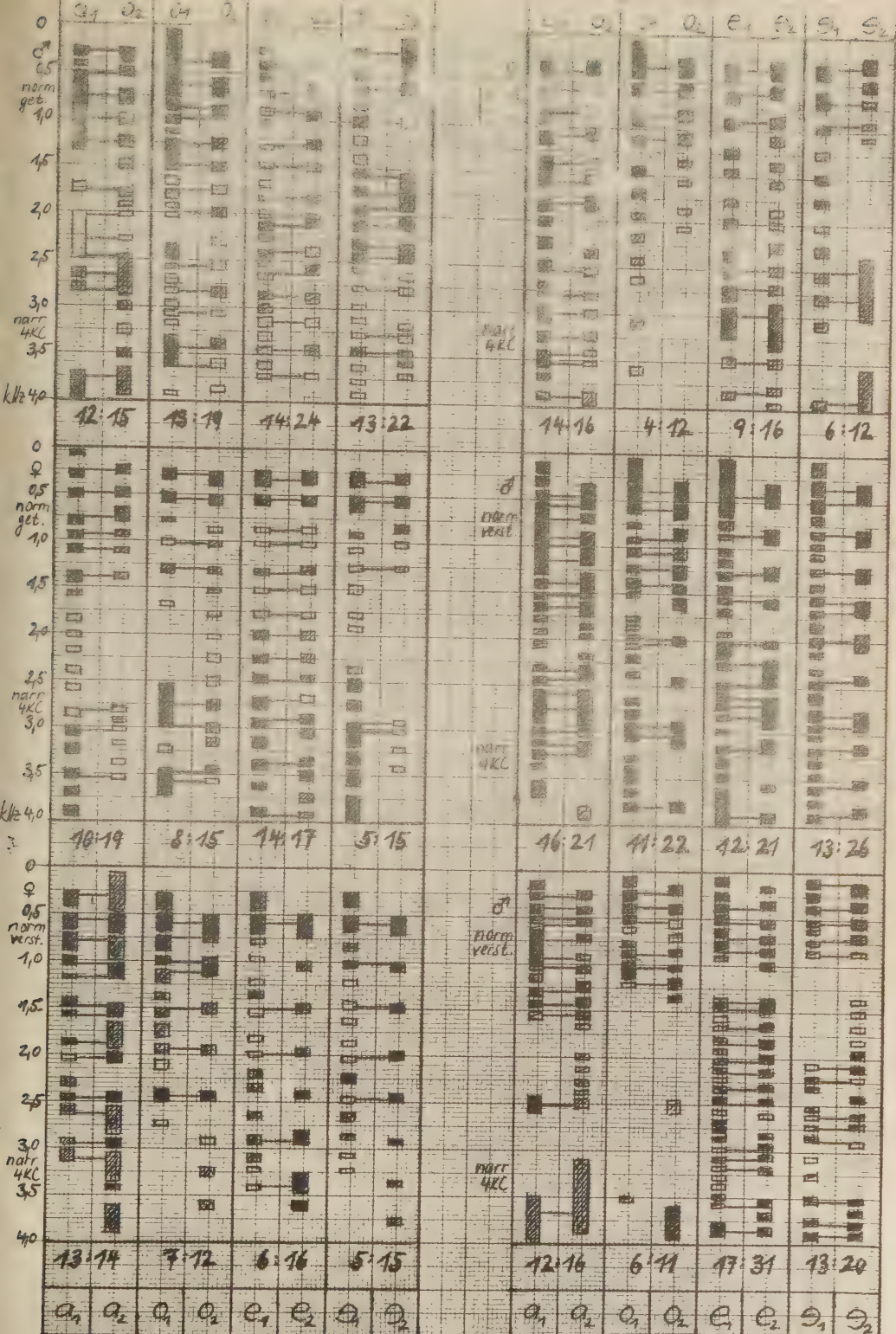
klz 8,0

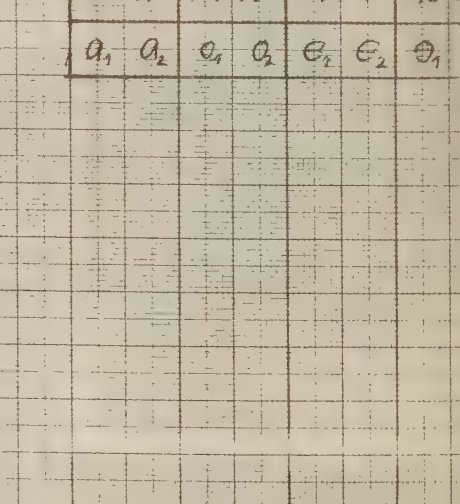
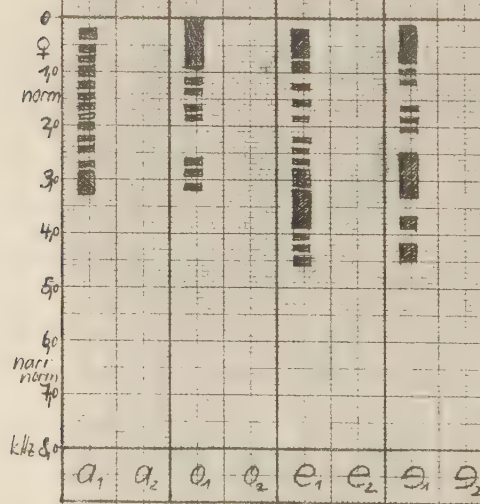
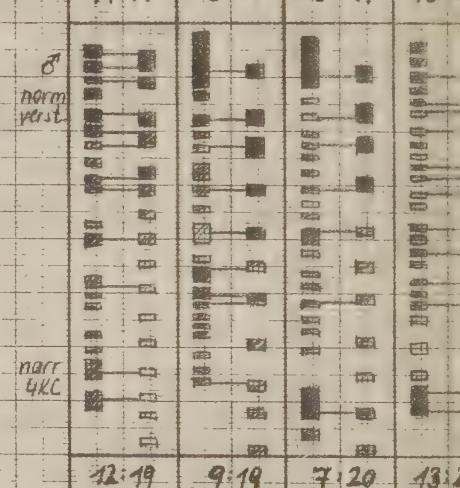
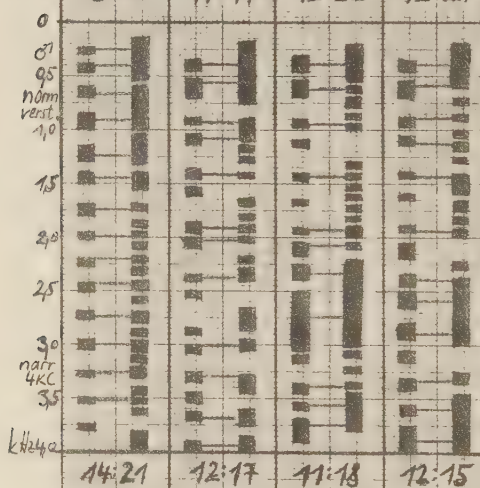
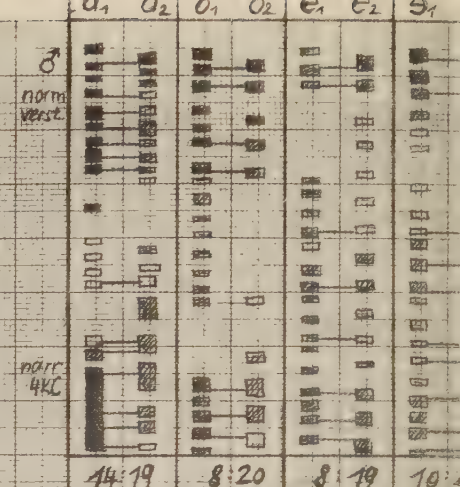
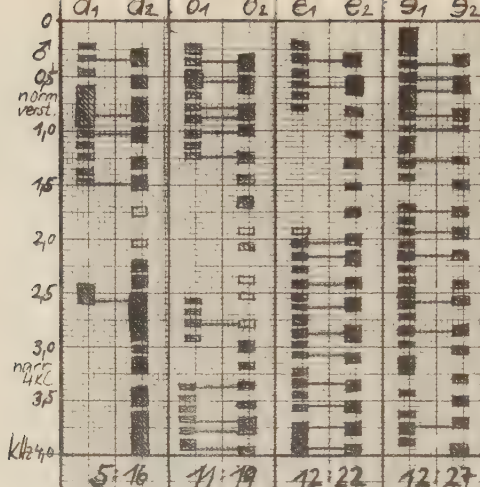


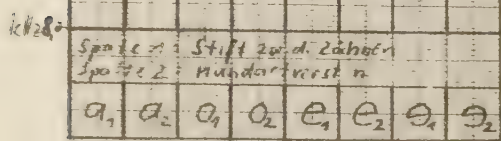
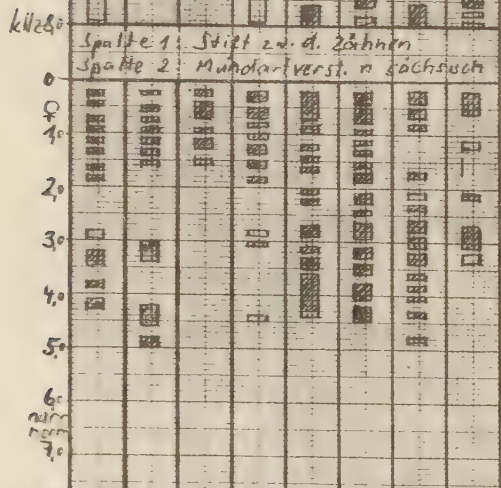
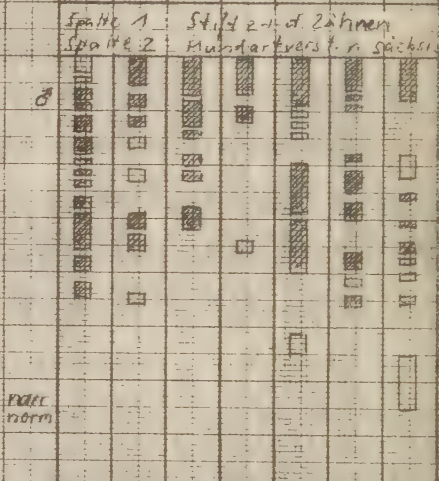
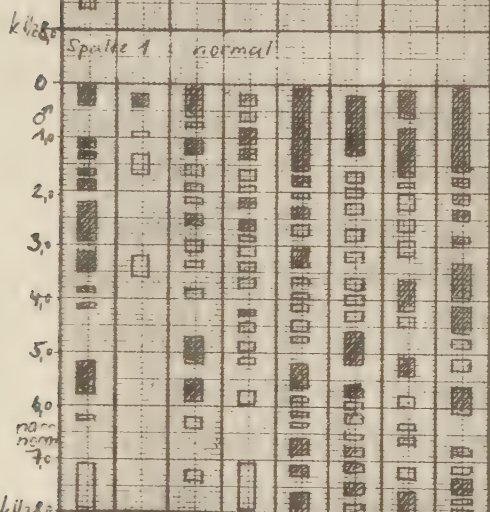
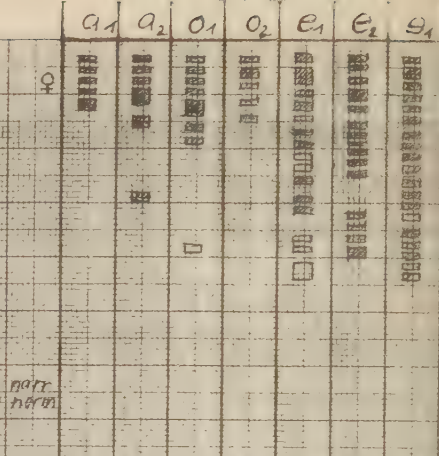
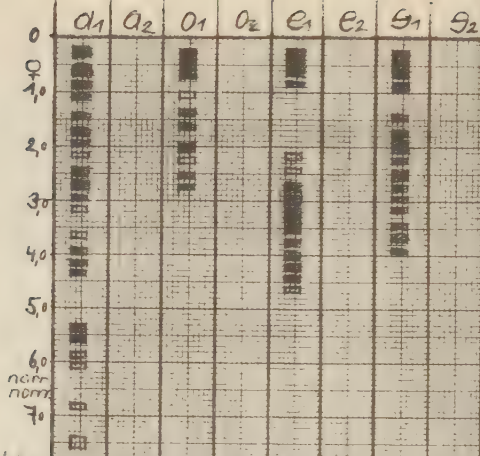
klz 8,0

Columns: a_1 , a_2 , e_1 , e_2









IV. Danksagungen

Mein besonderer Dank gilt Herrn Prof. Dr. E. Loebell, Leiter der Stimm- und Sprachabteilung der Universitäts-HNO-Klinik München, für die Überlassung dieser Arbeit und des Themas sowie für seine freundliche Unterstützung bei der Durchführung der Versuche und die Hilfe und Anleitung während der Fertigstellung der Arbeit.

Herrn Prof. Dr. H.H. Naumann, Leiter der Universitäts-HNO-Klinik München, danke ich für die Arbeitsmöglichkeit an seiner Klinik.

Großen Dank schulde ich Herrn Reg.-Chemiedirektor Dr. H. Habersbrunner, Leiter der kriminalistischen Abteilung des Bayerischen Landeskriminalamtes. Ich bedanke mich sehr für seine Hilfsbereitschaft und die Arbeitsmöglichkeit in seiner Abteilung.

Sehr herzlich danke ich Frau S. Hanel für ihre großen Hilfeleistungen bei der Herstellung der Sonagramme und Oszillographenaufnahmen sowie für Ihre fortgesetzte Hilfsbereitschaft in allen organisatorischen Fragen.

V. Lebenslauf .

Familienname:	Ludewig, geb. Roßmanith
Vorname:	Maria Magdalena
Geburtsdatum:	29.11.1944
Geburtsort:	Raase / CSR
Name des Vaters:	Dipl.-Chem. Gottfried Roßmanith
Name der Mutter:	Irene Roßmanith, geb. Jäckel
Volksschule:	Sept. 1951 Grundschule Großdeuben b. Leipzig Sept. 1957 Grundschule Krumpa Kr. Merseburg
Höhere Schule:	Sept. 1959 Erweiterte Oberschule Mücheln Kr. Merseburg Sept. 1961 Oberrealschule Bad Reichenhall Sept. 1963 Christoph-Scheiner-Oberrealschule Ingolstadt Juli 1965 dort Abitur
Hochschule:	Nov. 1965 Studium der Zahnheilkunde begonnen an der Ludwig-Maximilian-Universität München
Vorphysikum:	Sept./Okt. 1966 in München
Physikum:	März/April 1968 in München
Staatsexamen:	Sept. 1970 - April 1971 in München
Heirat:	Dez. 1968 mit Stud.-Ref. Hans-Ulrich Ludewig
Kind:	Marion Margarethe Ludewig

G LITERATURVERZEICHNIS

Arndt: zit. nach Ranke, O. und Lullies, H., Lehrbuch der Physiologie, Gehör, Stimme, Sprache. Springer Verlag, Berlin 1953.

Beckmann und Oeken: zit. nach Ranke, O. und Lullies, H., Lehrbuch der Physiologie, Gehör, Stimme, Sprache. Springer Verlag, Berlin 1953.

Böhme, Gerhard: Störungen der Sprache, der Stimme und des Gehörs durch frühkindliche Hirnschädigungen (Klinische und experimentell phonetische Untersuchungen), VEB Gustav Fischer Verlag, Jena 1966.

Böhme Gerhard: Stimm-, Sprach- und Hörstörungen als Folge frühkindlicher Hirnschädigungen. Folia phoniat. 15, 189-200, 1963

Böhme Gerhard: Klinische und experimentelle Untersuchungen über dysarthrische Sprachstörungen infolge infantiler Cerebralpareesen. HNO 15, 196 .

Berendes, J.: HNO-Heilkunde - ein kurzgefaßtes Handbuch in 3 Bänden. Band II, Teil 2, G. Thieme Verlag, Stuttgart, 1963.

Berendes, J. und Schilling, A.: Stimm- und Sprachstörungen. Lehmann, München 1962.

Dunn, H.K.: Methods of Measuring Vowel-Formant-Bandwidth. Journal of the Acoustical Society of America 33, 1737, 1961.

Dunn, J.L.: Airborne Measurements of the Acoustic Characteristics of a Sperm Whale. Journal of the Acoustical Society of America 46, Nr.4, 1963.

Essen, Otto von: Allgemeine und angewandte Phonetik. Akademie-Verlag, Berlin, 1957.

Essen, Otto von: Grundbegriffe der Phonetik. Berlin, 1962.

Essen, Otto von: Identifizierung von Sprechstimmen nach sonographischer Registrierung. Sonderdruck aus "Sprachheilpädagogische und hirnpathologische Probleme bei der Rehabilitation von Hirn- und Sprachgeschädigten"

(Tagungsbericht Köln 1966 -- Arbeitsgemeinschaft für Sprachheilpädagogik i.D.)

Ewing, W. u.a.: Transient Phenomena in Explosive Sound. Reprinted from "Proceedings 3rd International Congress on Acoustics". Elsevier Publishing Company, Amsterdam.

Ewing, M. u.a.: Dispersive Transients in Earthquake Signals. Reprinted from "Proceedings 3rd International Congress on Acoustics." Elsevier Publishing Company, Amsterdam.

Ey, W.: Vergleichende sprachanalytische Untersuchungen mittels der Sonographie an normalhörenden, hörrestigen und gehörlosen Schulkindern. Archiv für Ohren-Nasen-Kehlkopfheilkunde 180, 245-250, 1962.

Fährmann, Rudolf: Grundprobleme der Sprechstimmverstellung und Sprechstimmvergleichung. Ein Beitrag zur Praxis der Kriminalphonetik, insbesondere zur Methode der Sprechstimmexpertise. Arch. Kriminolog. 137, 25-32 und 91-102, 1966.

Fährmann, Rudolf: Die Deutung des Sprechausdrucks. Bonn 1960.

Fant, C.Gunnar M.: Modern Instruments and Methods for Acoustic Studies of Speech. Proceedings of the 8th International Congress of Linguists. Oslo 1958, 283.

Fant, C.G.M. u.a.: An Electrical Analog of the Vokal Tract. Journal of the Ac.Soc. of America, 25, Nr.4, 1953.

Flach, M. und Schwickardi, H.: Experimentelle Untersuchungen über den Einfluß der Schleimhutanästhesie auf den Stimmklang und über die Frage einer akustischen Funktion der recessus piriformes. HNO 15, 1967.

Fletcher, H.: zit. nach Luchsinger, R. und Arnold, G.E., Lehrbuch der Stimm- und Sprachheilkunde. Springer Verlag, Wien, 1959.

Habersbrunner, H.: Zur Personenfeststellung mittels Stimm- und Sprachanalyse, I. Teil. Arch. Kriminol. 1968.

Harris, C.M.: A Study of the Building Blocks in Speech. Journal of the Ac. Soc. of America 25, Nr.5, 962, 1953.

Handsel, Leon: Acoustic Analysis of Vowels in Deaf Children, by Means of the "Visible Speech" Apparatus. Folia phoniatrica 8, 237-240, 1956.

Helmholtz, H.v.: zit. nach Essen, O.v., Allgemeine und angewandte Phonetik, Akademie-Verlag, Berlin, 1957.

Hermann: zit. nach Essen, O.v., Allgemeine und angewandte Phonetik. Akademie-Verlag, Berlin, 1957.

Joos, Martin: Acoustic Phonetics, Supplement to Language. Journal of the Linguistic Society of America 24, 1948.

Kersta, L.G.: Amplitude Cross-Section Representation with the Sound-Spectrograph. Journal of the Ac. Soc. of America 20, 796-801, 1948.

Kersta, L.G.: Voiceprint Identification. Journal of the Ac.Soc. of America 34, 725, 1962.

Kersta, L.G.: Voiceprint Identification Infallibility. Paper for Presentation of the Ac. Soc. of America, 1978, 1962.

Koenig, W. und Ruppel, A.E.: Qualitative Amplitude Representation in Sound Spectrograms. Journal of the Ac. Soc. of America 20, 787-795, 1948.

Kreyszig, Erwin: Statistische Methoden und ihre Anwendung. Verlag Vandenhoeck und Ruprecht, Göttingen, 1967.

Loebell, E.: Frequenzzzerlegte Elektroglottographie als Beitrag zur Entstehungstheorie von Formanten. I. COMET Tagung, Barcelona, 1970.

Luchsinger, Richard und Arnold, G.E.: Lehrbuch der Stimm- und Sprachheilkunde. Springer-Verlag, Wien, 1959.

Mathews, M.V. u.a.: Pitch Synchronous Analysis of Voiced Sounds. Journal of the Ac. Soc. of America 33, 179, 1961.

Menon, K.M.N. u.a.: Acoustic Properties of Certain VCC Utterances. Journal of the Ac. Soc. of America 46, 447, 1969.

Potter, Ralph K., Kopp, G.A. und Green, H.C.: Visible Speech. D. van Nostrand Co., Princeton, New Jersey, 1947.

Potter, R.K. und Peterson, G.E.: The Presentation of Vowels and Their Movements. Journal of the Ac.Soc. of America 20, 528-535, 1948.

Peterson, G.E.: The Information Bearing Elements of Speech. Journal of the Ac. Soc. of America 24, 629, 1952.

Peterson, G.E. und Barnay, H.L.: Control Methods Used in a Study of the Vowels. Journal of the Ac. Soc. of America 24, 179-184, 1952.

Peterson, G.E. und Raisbeck, G.: The Measurement of Noise with the Sound Spectrograph. Journal of the Ac. Soc. of America 25, 1157-1162, 1953.

Prestigiacomo, A.J.: Amplitude Contour Display of Sound Spectrograph. Journal of the Ac. Soc. of America 34, 1684-1688, 1962.

Ranke, O. und Lullies, H.: Lehrbuch der Physiologie, Gehör, Stimme, Sprache. Springer Verlag, Berlin, 1953.

Sokolowsky, R.: Zur Charakteristik der Vokale. Zeitschrift für HNO 6, 557, 1923.

Stumpf, C.: zit. nach Essen, O.v., Allgemeine und angewandte Phonetik. Akademie-Verlag, Berlin, 1957.

Thielcke, Gerhard: Die Auswertung von Vogelstimmen nach Tonbandaufnahmen. Sonderdruck aus: Die Vogelwelt, 87. Jahrgang, Heft 1, 1966.

Todt, Dietmar: Gesang und gesangliche Korrespondenz der Amsel. Naturwissenschaften 57, 61-66, 1970.

Trendlenburg: zit. nach Ranke, O. und Lullies, H., Lehrbuch der Physiologie, Gehör, Stimme, Sprache. Springer Verlag, Berlin, 1953.

Trojan, F. und Kryspin-Exner, K.: The Decay of Articulation under the Influence of Alkohol and Paraldehyde. *Folia phoniatica* 4, 217, 1968.

Weiss, Mark R. und Harris, C.M.: Computer Technique for High-Speed Extraction of Speech Parameters. *Journal of the Ac. Soc. of America* 35, 207-214, 1963.

Zenker, W. und Zenker, A.: Über die Regelung der Stimmlippenspannung durch von außen angreifende Mechanismen. *Folia phoniatica* 12, 27, 1960.

